

5

1441 TTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA 1480

1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520

10

1521 CATTTCAGAAATAGAGGGTATATTGAAGTTCCAATTCACTTC 1560

1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600

15

1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640

1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680

20

1681 TCCTTGGATAATCTCCAATCCAGCGATTTCCGGTTACTTTG 1720

1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAAACATCGT 1760

25

1761 GGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800

30

1801 GACAGATTCGAGTTCATTCCAGTTACTGCAACACTCGAGG 1840

1841 CTGAGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGC 1880

35

1881 CCTCTTTACCTCCACCAATCAGCTTGGCTTGAAAATAAC 1920

1921 GTTACTGACTATCACATTGACCAAGTGTCCAACCTTGGTCA 1960

40

1961 CCTACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGA 2000

2001 ACTCTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGAC 2040

45

2041 GAGAGGAATCTCTTGCAAGACTCCAACCTTCAAAGACATCA 2080

2081 ACAGGCAGCCAGAACGTGGTTGGGGTGGGAAGCACCGGGAT 2120

50

2121 CACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC 2160

55

2161 GTCACCCCTCTCCGGAACCTTTCGACGAGTGCTACCCTACCT 2200

2201 ACTTGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTT 2240
 5
 2241 CACCAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA 2280
 10
 2281 GACCTTGAAATCTACTCGATCAGGTACAATGCCAAGCAGG 2320
 2321 AGACCGTGAATGTCCCAGGTACTGGTTCCTCTGGCCACT 2360
 15
 2361 TTCTGCCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCTAAC 2400
 2401 AGATGCGCTCCACACCTTGAGTGGAAATCCTGACTTGGACT 2440
 20
 2441 GCTCCTGCAGGGATGGCGAGAAGTGTGCCCACCATTCTCA 2480
 2481 TCACTTCTCCTTGGACATCGATGTGGGATGTACTGACCTG 2520
 25
 2521 AATGAGGACCTCGGAGTCTGGGTCATCTTCAAGATCAAGA 2560
 2561 CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCT 2600
 30
 2601 CGAAGAGAAACCATTTGGTCGGTGAAGCTCTCGCTCGTGTG 2640
 2641 AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAAC 2680
 35
 2681 TCGAATGGGAAACTAACATCGTTTACAAGGAGGCCAAAGA 2720
 40
 2721 GTCCGTGGATGCTTTGTTCGTGAACTCCCAATATGATCAG 2760
 2761 TTGCAAGCCGACACCAACATCGCCATGATCCACGCCGCAG 2800
 45
 2801 ACAAACGTGTGCACAGCATTCTGTGAGGCTTACTTGCCTGA 2840
 2841 GTTGTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAG 2880
 50
 2881 GAACTTGAGGGACGTATCTTTACCGCATTCTCCTTGTACG 2920
 55
 2921 ATGCCAGAAACGTCAATCAAGAACGGTGACTTCAACAATGG 2960

2961 CCTCAGCTGCTGGAATGTGAAAGGTCATGTGGACGTGGAG 3000
 3001 GAACAGAACAAATCAGCGTTCCGTCCTGGTTGTGCCTGAGT 3040
 3041 GGGAAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGG 3080
 3081 TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA 3120
 3121 TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACA 3160
 3161 ACACCGACGAGCTTAAGTTCTCCAAGTGCCTCGAGGAAGA 3200
 3201 AATCTATCCCAACAACACCGTTACTTGCAACGACTACACT 3240
 3241 GTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTA 3280
 3281 ACAGAGGTTACAACGAAGCTCCTTCCGTTCCCTGCTGACTA 3320
 3321 TGCCTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGA 3360
 3361 CGTGAGAACCCTTGCAGATTCAACAGAGGTTACAGGGACT 3400
 3401 ACACACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGA 3440
 3441 GTACTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGT 3480
 3481 GAAACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTC 3520
 3521 TCTTGATGGAGGAA 3534,

H. einem Struktur-Gen, das für ein insektizides Protein von *B.t.f.* codiert, mit der Sequenz:

1 ATGACTGCAGACAACAACACCGAAGCCCTCGACAGTTCTA 40
 41 CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG 80

5

81 AGACCTCTTGGGCGTGGTTGGATTTCCTTCGGTGGAGCC 120

121 CTCGTGAGCTTCTATACAACTTTCTCAACACCATTGTC 160

10

161 CAAGCGAGGACCCTTGGAAAGCATTCATGGAGCAAGTTGA 200

201 AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC 240

15

241 AAGGCTTTGGCAGAACTCCAGGGCCTTCAGAACAATGTGG 280

281 AGGACTACGTGAGTGCATTGTCCAGCTGGCAGAAGAACCC 320

20

321 TGTTAGCTCCAGAAATCCTCAGCCAAAGGTAGGATCAGA 360

361 GAGTTGTTCTCTCAAGCCGAATCCCACTTCAGAAATTCCA 400

25

401 TGCCTAGCTTTGCTATCTCCGGTTACGAGGTTCTTTTCCT 440

30

441 CACTACCTATGCTCAAGCTGCCAACACCCACTTGTTTCTC 480

481 CTTAAGGACGCTCAAATCTATGGAGAAGAGTGGGGATACG 520

35

521 AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTTAA 560

561 GCTCACCCAAGAGTACACTGACCATTGCGTGAAATGGTAT 600

40

601 AACGTTGGTCTCGATAAGCTCAGAGGCTCTTCCTACGAGT 640

641 CTTGGGTGAACTTCAACAGATACAGGAGAGAGATGACCTT 680

45

681 GACTGTGCTCGATCTTATCGCACTCTTCCCTTGTACGAT 720

50

721 GTGAGACTCTACCCAAAGGAAGTGAAACTGAGCTTACCA 760

761 GAGACGTGCTCACTGACCCTATTGTGCGGAGTCAACAACCT 800

55

801 TAGGGGTTATGGAACTACCTTCAGCAATATCGAAAACCTAC 840

5

841 ATTAGGAAACCACATCTCTTCGACTATCTTCACAGAATTC 880

881 AATTCCACACAAGGTTTCAACCAGGATACTATGGTAACGA 920

10

921 CTCCTTCAACTATTGGTCCGGTAACTATGTTTCCACCAGA 960

961 CCAAGCATTGGATCTAATGACATCATCACATCTCCCTTCT 1000

15

1001 ATGGTAACAAGTCCAGTGAACCTGTGCAGAACCTTGAGTT 1040

1041 CAACGGCGAGAAAGTCTATAGAGCCGTGCGAAACACCAAT 1080

20

1081 CTCGCTGTGTGGCCATCCGCAGTTTACTCAGGCGTCACAA 1120

1121 AGGTGGAGTTTGTAGTCAGTATAACGATCAGACCGATGAGGC 1160

25

1161 CAGCACCCAGACTTACGACTCCAAACGTAACGTTGGCGCA 1200

30

1201 GTCTCTTGGGATTCTATCGACCAATTGCCTCCAGAAACCA 1240

1241 CAGACGAACCATTGGAGAAGGGCTACAGCCACCAACTTAA 1280

35

1281 CTATGTGATGTGCTTCTTGATGCAAGGTTCCAGAGGGACC 1320

1321 ATTCCAGTGTTGACCTGGACACACAAGTCCGTGGACTTCT 1360

40

1361 TCAACATGATCGATAGCAAGAAGATCACTCAACTTCCCTT 1400

1401 GGTGAAAGCCTACAAGCTGCAATCTGGTGCTTCCGTTGTC 1440

45

1441 GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA 1480

1481 CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA 1520

50

1521 TGTGTCTTACTCTCAGAAGTACAGGGCACGTATTCATTAC 1560

55

1561 GCATCTACCAGCCAGATCACCTTCACACTCAGCTTGGATG 1600

1601 GAGCACCCTTCAACCAGTATTACTTTGACAAGACCATCAA 1640
 1641 CAAAGGTGACACTCTCACATAACAATAGCTTCAACTTGGCA 1680
 1681 AGTTTCAGCACACCATTGAACTCTCAGGCAACAATCTTC 1720
 1721 AGATCGGCGTCACCGGTCTCAGCGCCGGAGACAAAGTCTA 1760
 1761 CATCGACAAGATTGAGTTCATCCCAGTGAAC 1791

I. einem Struktur-Gen, das für ein insektizides Protein von *B. t. entomocidus* codiert, mit der Sequenz:

1 ATGGAGGAGAACAACCAAAACCAATGCATTCCATACAACT 40
 41 GCTTGAGTAACCCAGAAGAGGTATTGCTTGATGGAGAACG 80
 81 CATTTCAACCGGTAACCTCTCCATCGACATCTCCTTGTC 120
 121 TTGGTCCAGTTTCTGGTCAGCAACTTCGTGCCAGGTGGTG 160
 161 GGTTCCCTTGTCGGACTAATTGACTTCGTCTGGGGTATCGT 200
 201 TGGTCCATCTCAATGGGATGCATTCTGGTGCAAATTGAG 240
 241 CAGTTGATCAACGAGAGGATCGCTGAGTTCGCCAGGAACG 280
 281 CTGCCATCGCTAACTTGAAGGATTGGGCAATAACTTCAA 320
 321 CATCTATGTGGAGGCCTTCAAAGAGTGGGAAGAGGACCCT 360
 361 AACAAACCAGAGACCCGCACTAGGGTGATCGACAGATTCA 400
 401 GAATCTTGGACGGCCTCTTGGAGAGAGATATCCCATCCTT 440
 441 CAGAATCTCTGGCTTCGAAGTTCCTCTCTTGTCCGTGTAC 480

5

481 GCTCAAGCAGCTAATCTTCACCTCGCTATCCTTCGAGACA 520

521 GTGTCATCTTTGGGGAAAGGTGGGGATTGACCACTATCAA 560

10

561 CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC 600

601 GAGTACGCCGACCACTGTGCTAACACCTACAACCGTGGCT 640

15

641 TGAACAATCTCCCTAAGTCTACTTATCAAGATTGGATTAC 680

681 CTACAACAGGTTGAGGAGAGACTTGACCCTCACAGTTTTG 720

20

721 GACATTGCAGCTTTCTTCCCGAACTATGACAACAGGAGAT 760

761 ACCCTATCCAACCACTGGGTCAACTTACCAGAGAAGTCTA 800

25

801 TACTGACCCACTTATCAACTTCAACCCTCAGTTGCAAAGT 840

30

841 GTCGCCCAACTTCCCACATTCAACGTCATGGAGTCCAGCC 880

881 GTATCAGGAACCCACACTTGTTTGACATCTTGAACAACCT 920

35

921 TACTATCTTCACCGATTGGTTTCAGCGTTGGGCGTAACTTC 960

961 TATTGGGGTGGACACAGGGTCATCTCCTCTCTTATTGGAG 1000

40

1001 GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGCAA 1040

1041 CCAGGAGCCACCACGTAGTTTCACCTTCAACGGTCCAGTC 1080

45

1081 TTCAGAACCTTGTCTAACCTACCTTGAGATTGCTCCAGC 1120

50

1121 AACCTTGGCCAGCTCCACCTTTCAACCTTAGAGGTGTTGA 1160

1161 GGGCGTTGAGTTCTCTACTCCTACCAACTCCTTCACTTAC 1200

55

1201 AGAGGTAGAGGAACCGTTGATTCTTGACCGAACTCCAC 1240

5	1241	CAGAGGACAATAGCGTGCCACCCAGGGAAGGCTACTCCCA	1280
	1281	CAGGTTGTGCCACGCAACCTTCGTGCAGCGTTCCGGA	1320
10	1321	CCATTCCTCACTACAGGAGTTGTGTTCTCATGGACTGATC	1360
	1361	GTAGTGCTACTCTCACTAATACCATTGATCCCGAGAGGAT	1400
15	1401	CAATCAAATCCCATTGGTCAAGGGTTTCCGTGTGTGGGGA	1440
	1441	GGAACCTTCTGTCAATCACAGGACCAGGCTTCACAGGAGGTG	1480
20	1481	ATATTCCTTAGAAGAAACACTTTTGGCGACTTTGTGAGCCT	1520
	1521	CCAAGTTAACATCAACTCTCCAATTACTCAAAGATATCGT	1560
25	1561	CTCAGGTTTCGTTACGCATCTTCCCGTGACGCTAGAGTCA	1600
	1601	TCGTGCTCACCAGGAGCAGCTTCTACCGGTGTGGTGGACA	1640
30	1641	AGTCTCCGTGAACATGCCACTCCAGAAGACTATGGAGATC	1680
	1681	GGCGAGAACTTGACATCCAGGACCTTCAGATACACCGACT	1720
35	1721	TCTCTAACCCCTTTTCACTTCCGTGCCAACCCCTGACATCAT	1760
40	1761	TGGCATTAGCGAACAACCTCTCTTTGGAGCTGGTAGCATC	1800
	1801	TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC	1840
45	1841	TTGCCGACGCTACCTTCGAGGCTGAGTCTGACCTTGAGAG	1880
	1881	AGCCCAGAAGGCTGTGAACGCCCTCTTTACCTCCTCTAAT	1920
50	1921	CAGATTGGCTTGAAAACTGACGTTACTGACTATCACATTG	1960
55	1961	ACCAAGTGTCCAACCTTGGTCGACTGCCTTAGCGATGAGTT	2000

5

2001 CTGCCTCGACGAGAAGCGTGAAGTCTCCGAGAAAGTTAAA 2040

2041 CACGCCAAGCGTCTCAGCGACGAGAGGAATCTCTTGCAAG 2080

10

2081 ACCCCAAGTTCAGAGGCATCAACAGGCAGCCAGACCGTGG 2120

2121 TTGGAGAGGAAGCACCGACATCACCATCCAAGGAGGCGAC 2160

15

2161 GATGTGTTCAAGGAGAACTACGTCACCCTCCCAGGAACTG 2200

2201 TGGACGAGTGCTACCCCTACCTACTTGTACCAGAAGATCGA 2240

20

2241 TGAGTCCAAACTCAAAGCCTACACCAGGTATGAACTTAGA 2280

25

2281 GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA 2320

2321 TCAGGTACAATGCCAAGCACGAGATCGTGAATGTCCCAGG 2360

30

2361 TACTGGTTCCCTCTGGCCACTTTCTGCCCAAATGCCCAT 2400

2401 GGGAAAGTGTGGAGAGCCTAACAGATGCGCTCCACACCTTG 2440

35

2441 AGTGGAATCCTGACTTGGACTGCTCCTGCAGGGATGGCGA 2480

2481 GAAGTGTGCCACCATTTCTCATCACTTCACCTTGGACATC 2520

40

2521 GATGTGGGATGTACTGACCTGAATGAGGACCTCGGAGTCT 2560

2561 GGGTCATCTTCAAGATCAAGACCCAAGACGGACACGCAAG 2600

45

2601 ACTTGGCAACCTTGAGTTTCTCGAAGAGAAACCATTGCTC 2640

50

2641 GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAAGT 2680

2681 GGAGGGACAAACGTGAGAACTCCAACCTCGAGACTAACAT 2720

55

2721 CGTTTACAAGGAGGCCAAAGAGTCCGTGGATGCTTTGTTC 2760

2761	GTGAACTCCCAATATGATAGGTTGCAAGTGGACACCAACA	2800
2801	TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT	2840
2841	TCGTGAGGCTTACTTGCCTGAGTTGTCCGTGATCCCTGGT	2880
2881	GTGAACGCTGCCATCTTCGAGGAACCTGAGGGACGTATCT	2920
2921	TTACCGCATACTCCTTGTACGATGCCAGAAACGTCATCAA	2960
2961	GAACGGTGACTTCAACAATGGCCTCTTGTGCTGGAATGTG	3000
3001	AAAGGTCATGTGGACGTGGAGGAACAGAACAATCACCGTT	3040
3041	CCGTCCTGGTTATCCCTGAGTGGGAAGCTGAAGTGTCCCA	3080
3081	AGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGT	3120
3121	GTGACCGCTTACAAGGAGGGATACGGTGAGGGTTGCGTGA	3160
3161	CCATCCACGAGATCGAGGACAACACCGACGAGCTTAAGTT	3200
3201	CTCCAAGTGCCTCGAGGAAGAAGTCTATCCCAACAACACC	3240
3241	GTTACTTGCAACAACACTACACTGGGACCCAGGAAGAGTACG	3280
3281	AAGGTACCTACACTAGCCGTAACCAAGGTTACGACGAAGC	3320
3321	TTACGGAAACAATCCTTCCGTTCTGCTGACTATGCCTCC	3360
3361	GTGTACGAGGAGAAATCCTACACAGATGGCAGACGTGAGA	3400
3401	ACCCTTGCGAGTCCAACAGAGGTTACGGTGACTACACACC	3440
3441	ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT	3480
3481	CCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAAACCG	3520

3521 AGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCTTGAT 3560

3561 GGAGGAA 3567.

J. einem Struktur-Gen, das für ein insektizides P2-Protein codiert, mit der Sequenz:

1 ATGGACAACAACGTCTTGAACCTCTGGTAGAACAACCATCT 40
 41 GCGACGCATACAACGTCTGGCTCACGATCCATTTCAGCTT 80
 81 CGAACACAAGAGCCTCGACACTATTCAGAAGGAGTGGATG 120
 121 GAATGGAAACGTACTGACCACTCTCTCTACGTGCGACCTG 160
 161 TGGTTGGAACAGTGTCCAGCTTCCTTCTCAAGAAGGTCGG 200
 201 CTCTCTCATCGGAAAACGTATCTTGTCCGAACTCTGGGGT 240
 241 ATCATCTTTCCATCTGGGTCCACTAATCTCATGCAAGACA 280
 281 TCTTGAGGGAGACCGAACAGTTTCTCAACCAGCGTCTCAA 320
 321 CACTGATACCTTGGCTAGAGTCAACGCTGAGTTGATCGGT 360
 361 CTCCAAGCAAACATTCGTGAGTTCAACCAGCAAGTGGACA 400
 401 ACTTCTTGAATCCAACCTCAGAATCCTGTGCCTCTTTCCAT 440
 441 CACTTCTTCCGTGAACACTATGCAGCAACTCTTCCTCAAC 480
 481 AGATTGCCTCAGTTTCAGATTCAAGGCTACCAGTTGCTCC 520
 521 TTCTTCCACTCTTTGCTCAGGCTGCCAACATGCACTTGTC 560
 561 CTTCATACGTGACGTGATCCTCAACGCTGACGAATGGGGA 600

5

601 ATCTCTGCAGCCACTCTTAGGACATACAGAGACTACTTGA 640

641 GGAACTACACTCGTGATTACTCCAACCTATTGCATCAACAC 680

10

681 TTATCAGACTGCCTTTCGTGGACTCAATACTAGGCTTCAC 720

721 GACATGCTTGAGTTCAGGACCTACATGTTCTTAACGTGT 760

15

761 TTGAGTACGTCAGCATTGAGTCTCTTCAAGTACCAGAG 800

801 CTTGATGGTGTCTCTGGAGCCAATCTCTACGCCTCTGGC 840

20

841 AGTGGACCACAGCAAACTCAGAGCTTCACAGCTCAGAACT 880

25

881 GGCCATTCTTGATAGCTTGTTCCAAGTCAACTCCAATA 920

921 CATTCTCAGTGGTATCTCTGGGACCAGACTCTCCATAACC 960

30

961 TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC 1000

1001 ATAGCCTTAACTCTGCCAGAGTGAACCTACAGTGGAGGTGT 1040

35

1041 CAGCTCTGGATTGATTGGTGCAACTAACTTGAACCACAAC 1080

1081 TTCAATTGCTCCACCGTCTTGCCACCTCTGAGCACACCGT 1120

40

1121 TTGTGAGGTCCTGGCTTGACAGCGGTACTGATCGCGAAGG 1160

1161 AGTTGCTACCTCTACAACTGGCAAACCGAGTCCTTCCAA 1200

45

1201 ACCACTCTTAGCCTTCGGTGTGGAGCTTCTCTGCACGTG 1240

50

1241 GGAATTCAAACCTACTTTCCAGACTACTTCATTAGGAACAT 1280

1281 CTCTGGTGTTCCTCTCGTCATCAGGAATGAAGACCTCACC 1320

55

1321 CGTCCACTTCATTACAACCAGATTAGGAACATCGAGTCTC 1360

1361 CATCCGGTACTCCAGGAGGTGCAAGAGCTTACCTCGTGTC 1400
 1401 TGTCCATAACAGGAAGAACAACATCTACGCTGCCAACGAG 1440
 1441 AATGGCACCATGATTACCTTGCACCAGAAGATTACACTG 1480
 1481 GATTCACCATCTCTCCAATCCATGCTACCCAAGTGAACAA 1520
 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560
 1561 GGTGACTCCTTGAGGTTGAGCAATCCAACACTACCGCTA 1600
 1601 GGTACACTTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640
 1641 CTTGAGAGTTAGCTCCATTGGTAACTCCACCATCCGTGTT 1680
 1681 ACCATCAACGGACGTGTTTACACAGTCTCTAATGTGAACA 1720
 1721 CTACAACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760
 1761 ATTCAGCGACATCAACATTGGCAACATCGTGGCCTCTGAC 1800
 1801 AACACTAACGTTACTTTGGACATCAATGTGACCCTCAATT 1840
 1841 CTGGAACCTCATTGATCTCATGAACATCATGTTGTGCC 1880
 1881 AACTAACCTCCCTCCATTGTAC 1902

oder

K. einer Struktur-Gen-Sequenz, die für ein Fusionsprotein codiert, das die N-terminalen 610 Aminosäuren von B.t.k. HD-1 und die C-terminalen 567 Aminosäuren von B.t.k. HD-73 aufweist, welches Gen die Sequenz hat:

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40

41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 5 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 10 161 CTGGGTTCTGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200
 201 CTTTGGTCCATCTCAATGGGATGCATTCTGGTGCAAATT 240
 15 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 20 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360
 25 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 30 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG 480
 35 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560
 40 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
 601 GGAAACTACACCGACCAAGCTGTTCGTTGGTACAACACTG 640
 45 641 GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT 680
 681 TAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT 720
 50 721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760
 55 761 CCTACCCTATCCGTACAGTGTCCCACTTACCAGAGAAAT 800

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
	841	CGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
10	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
15	961	ATCATGGCCTCTCCAGTTGGATTTCAGCGGGCCCCGAGTTTA	1000
	1001	CCTTTCCTCTCTATGGAACCTATGGGAAACGCCGCTCCACA	1040
20	1041	ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA	1080
25	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG	1120
	1121	GSTATCAACAACCAGCAACTTCCGTTCTTGACGGAACAGA	1160
30	1161	GTTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
	1201	TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCC	1240
35	1241	CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC	1280
40	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC	1320
	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
45	1361	CATGGATTCATCGTAGTGCTGAGTTCAACAATATCATTCC	1400
	1401	TTCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT	1440
50	1441	AACCTTGGATCTGGAACCTTCTGTCGTGAAAGGACCAGGCT	1480
55	1481	TCACAGGAGGTGATATTCTTAGAAGAACTTCTCCTGGCCA	1520

5

1521 GATTAGCACCCCTCAGAGTTAACATCACTGCACCACTTTCT 1560

1561 CAAAGATATCGTGTTCAGGATTCGTTACGCATCTACCACTA 1600

10

1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640

1641 TCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680

15

1681 TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTC 1720

1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTTCAACCCTTAG 1760

20

1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGACATTGAC 1800

1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCCTCGAGGCTG 1840

25

1841 AGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGCCCT 1880

30

1881 CTTTACCTCCACCAATCAGCTTGGCTTGAAAACCTAACGTT 1920

1921 ACTGACTATCACATTGACCAAGTGTCCAACCTGGTCACT 1960

35

1961 ACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGAAC 2000

2001 CTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGACGAG 2040

40

2041 AGGAATCTCTTGCAAGACTCCAACCTCAAAGACATCAACA 2080

2081 GGCAGCCAGAACGTGGTTGGGGTGAAGCACCGGGATCAC 2120

45

2121 CATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTACGTC 2160

2161 ACCCTCTCCGGAACTTTCGACGAGTGCTACCCTACCTACT 2200

50

2201 TGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTTCAC 2240

55

2241 CAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAAGAC 2280

5

2281 CTTGAAATCTACTCGATCAGGTACAATGCCAAGCACGAGA 2320

2321 CCGTGAATGTCCCAGGTACTGGTCCCTCTGGCCACTTTC 2360

10

2361 TGCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCTAACAGA 2400

2401 TGCGCTCCACACCTTGAGTGGGAATCCTGACTTGGACTGCT 2440

15

2441 CCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCATCA 2480

2481 CTTCTCCTTGGACATCGATGTGGGATGTACTGACCTGAAT 2520

20

2521 GAGGACCTCGGAGTCTGGGTCATCTTCAAGATCAAGACCC 2560

2561 AAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCTCGA 2600

25

2601 AGAGAAACCATTGGTTCGGTGAAGCTCTCGCTCGTGTGAAG 2640

2641 AGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAACTCG 2680

30

2681 AATGGGAAACTAACATCGTTTACAAGGAGGCCAAAGAGTC 2720

2721 CGTGGATGCTTTGTTTCGTGAACCTCCCAATATGATCAGTTG 2760

35

2761 CAAGCCGACACCAACATCGCCATGATCCACGCCGCAGACA 2800

2801 AACGTGTGCACAGCATTTCGTGAGGCTTACTTGCCTGAGTT 2840

40

2841 GTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAGGAA 2880

45

2881 CTTGAGGGACGTATCTTTACCGCATTCTCCTTGTACGATG 2920

2921 CCAGAAACGTCATCAAGAACGGTGAAGTTCAACAATGGCCT 2960

50

2961 CAGCTGCTGGAATGTGAAAGGTCATGTGGACGTGGAGGAA 3000

55

3001 CAGAACAATCAGCGTTCCGTCCTGGTTGTGCCTGAGTGGG 3040

3041 AAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTAG 3080
 3081 AGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGATAC 3120
 3121 GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAACA 3160
 3161 CCGACGAGCTTAAGTTCTCCAAGTGGTTCGAGGAAGAAAT 3200
 3201 CTATCCCAACAACACCGTTACTTGCAACGACTACACTGTG 3240
 3241 AATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACA 3280
 3281 GAGGTTACAACGAAGCTCCTTCCGTTCTGCTGACTATGC 3320
 3321 CTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGACGT 3360
 3361 GAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACA 3400
 3401 CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA 3440
 3441 CTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA 3480
 3481 ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCT 3520
 3521 TGATGGAGGAA 3531.

Revendications

1. Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :

- a) l'identification de régions à l'intérieur de ladite séquence comprenant plus de quatre nucléotides consécutifs d'adénine ou de thymine,
- b) la modification des régions de l'étape a) qui comportent deux ou plusieurs signaux de polyadénylation à l'intérieur d'une séquence de dix bases afin d'éliminer lesdits signaux tout en conservant une séquence de gène qui code ladite protéine, et
- c) la modification des régions de 15 à 30 bases entourant les régions de l'étape a) afin d'éliminer les signaux majeurs de polyadénylation de plantes, les séquences consécutives contenant plus d'un signal mineur de polyadénylation et les séquences consécutives contenant plus d'une séquence ATTTA tout en conservant une séquence de gène qui code ladite protéine.

2. Procédé de modification d'une séquence de gène de structure du type sauvage qui code une protéine insecticide de *Bacillus thuringiensis* afin d'activer l'expression de ladite protéine chez des plantes qui comprend :

- a) l'élimination des signaux de polyadénylation contenus dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine, et
b) l'élimination des séquences ATTTA contenues dans ledit gène de type sauvage tout en conservant une séquence qui code ladite protéine.

3. Procédé selon la revendication 2, comprenant en outre l'élimination des séquences autoccomplémentaires et le remplacement de telles séquences par de l'ADN non autoccomplémentaire comprenant des codons préférés des plantes tout en conservant une séquence de gène de structure codant ladite protéine.

4. Procédé selon les revendications 1 à 3, comprenant en outre l'utilisation des séquences préférées des plantes au cours de l'élimination des signaux de polyadénylation et des séquences ATTTA.

5. Procédé selon les revendications 1 à 3, dans lequel les signaux de polyadénylation des plantes sont choisis parmi le groupe constitué de AATAAA, AATAAT, AACCAA, ATATAA, AATCAA, ATACTA, ATAAAA, ATGAAA, AAGCAT, ATTAAT, ATACAT, AAAATA, ATTAAA, AATTAA, AATACA et CATAAA.

6. Procédé destiné à améliorer l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans les cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont une partie au moins est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et ladite séquence de structure codante ne contient pas plus de 5 nucléotides consécutifs constitués de restes soit adénine, soit thymine.

7. Procédé d'amélioration de l'expression d'un gène hétérologue chez des plantes dans lequel ledit gène comprend un gène chimère modifié comprenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une protéine de *Bacillus thuringiensis*, dans lequel ledit procédé comprend la modification de ladite séquence de structure codante de sorte que ladite séquence comporte une séquence d'ADN qui diffère de la séquence d'ADN qui apparaît dans la nature codant ladite protéine de *Bacillus thuringiensis* et présente les caractéristiques suivantes :

ladite séquence de structure codante comporte une région qui est complémentaire de la séquence suivante :

GGCTTGATTCCCTAGCGAACTCTTCGATTCTCTGGTTGATGAGCTGTTTC

1 5 10 15 20 25 30 35 40 45

ladite région dans ladite séquence codante ayant éliminé 2 séquences AACCAA et 1 séquence AATTAA.

8. Procédé selon la revendication 7, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée de *Bacillus thuringiensis* kurstakis HD-1.

9. Procédé selon la revendication 7 ou 8, dans lequel la plante est un plan de tabac.

10. Gène chimère modifié contenant un promoteur qui agit dans des cellules végétales liées de façon fonctionnelle à une séquence de structure codante et à une région 3' non traduite contenant un signal de polyadénylation qui agit chez des plantes pour provoquer l'addition de nucléotides de polyadénylate sur l'extrémité 3' de l'ARN, dans lequel ladite séquence de structure codante code une protéine insecticide dont au moins une partie est dérivée d'une

protéine de *Bacillus thuringiensis*, dans lequel ladite séquence de structure codante comporte une séquence d'ADN qui diffère de la séquence d'ADN apparaissant dans la nature codant ladite protéine de *Bacillus thuringiensis* et est choisie à partir de :

5 A. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

10

15

20

25

30

35

40

45

50

55

5

1 ATGGCTATAGAACTGGTTACACCCCAATCGATATTTCCCT 40

41 TGTCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGG 80

10

81 TGCTGGATTTGTGTTAGGACTAGTTGATATTATCTGGGGA 120

121 ATTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAA 160

15

161 TTGAACAGCTCATCAACCAGAGAATCGAAGAGTTCGCTAG 200

201 GAATCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTT 240

20

241 TATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAG 280

25

281 ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA 320

321 ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCCT 360

30

361 CTTTTTGCGAGTTCAAAATTATCAAGTTCCTCTCCTCTCCG 400

401 TGTACGTTCAAGCTGCCAACCTCCACCTCTCAGTTTTGAG 440

35

441 AGATGTTTCAGTGTTTGGACAAAGGTGGGGATTGATGCC 480

40

481 GCGACTATCAATAGTCGTTATAATGATTAACTAGGCTTA 520

521 TTGGCAACTATACAGATCATGCTGTACGCTGGTACAATAC 560

45

561 GGGATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGG 600

601 ATCAGGTACAACCAGTTCAGAAGAGAGCTTACACTAACTG 640

50

641 TATTAGATATCGTTTCTCTATTTCCGAAGTATGATAGTAG 680

55

681 AACGTATCCAATTCGAACAGTTTCCCAATTAACAAGAGAA 720

5 721 ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT 760

761 TTCGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAG 800

10 801 TCCACATTTGATGGATATACTTAATAGTATAACCATCTAT 840

841 ACGGATGCTCATAGAGGAGAATACTACTGGTCCGGTCACC 880

15 881 AGATCATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATT 920

20 921 CACTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCA 960

961 CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATA 1000

25 1001 GAACATTATCGTCCACCTTATATAGAAGACCTTTTAAACAT 1040

1041 CGGGATCAACAACCAACAACCTATCTGTTCTTGACGGGACA 1080

30 1081 GAATTGCTTATGGAACCTCCTCAAATTGCCATCCGCTG 1120

35

40

45

50

55

1121 TATACAGAAAAAGCGGAACGGTAGATTGCGCTGGATGAAAT 1160
 5
 1161 ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT 1200
 10
 1201 AGTCATCGATTAAGCCATGTTTCAATGTTTCGTTTCAGGCT 1240
 1241 TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT 1280
 15
 1281 CTCTTGGATACATCGTAGTGCTGAGTTCAACAACATCATC 1320
 1321 CCTTCATCACAAATCACCCAAATCCCACTCACCAAGTCTA 1360
 20
 1361 CTAATCTTGGCTCTGGAACCTTCTGTCGTTAAAGGACCAGG 1400
 25
 1401 ATTTACAGGAGGAGATATTCTTCTGAAGAACTTCACCTGGC 1440
 1441 CAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTTAT 1480
 30
 1481 CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCAC 1520
 35
 1521 AAACCTTCAGTTCCACACATCAATTGACGGAAGACCTATT 1560
 1561 AATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA 1600
 40
 1601 ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTAC 1640
 1641 TCCGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTA 1680
 45
 1681 AGTGCTCATGTCTTCAATTGAGGCAATGAAGTTTATATAG 1720
 50
 1721 ATCGAATTGAATTTGTTCCGGCA 1743.

B. Un gène de structure qui code une protéine insecticide de *B.t.k.* HD-73 comportant la séquence :

5

1 ATGGCCATTGAAACCGGTTACACTCCCATCGACATCTCCT 40

41 TGTCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGG 80

10

81 TGCTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGT 120

121 ATCTTTGGTCCATCTCAATGGGATGCATTCTTGGTGCAA 160

15

161 TTGAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAG 200

201 GAACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTC 240

20

241 TACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCG 280

25

281 ATCCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCA 320

321 ATTCAACGACATGAACAGCGCCTTGACCACAGCTATCCCA 360

30

361 TTGTTGCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCG 400

35

401 TGTACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCG 440

40

45

50

55

441 AGACGTTAGCGTGTGTTGGGCAAAGGTGGGGATTTCGATGCT 480
 481 GCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGA 520
 521 TTGGAAACTACACCGACCACGCTGTTTCGTTGGTACAACAC 560
 561 TGGCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGG 600
 601 ATTAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAG 640
 641 TTTTGGACATTGTGTCTCTCTCCCGAACTATGACTCCAG 680
 681 AACCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAA 720
 721 ATCTATACTAACCCAGTTCCTTGAGAACTTCGACGGTAGCT 760
 761 TCCGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAG 800
 801 CCCACACTTGATGGACATCTTGAACAGCATAACTATCTAC 840
 841 ACCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACC 880
 881 AGATCATGGCCTCTCCAGTTGGATTTCAGCGGGCCCGAGTT 920
 921 TACCTTTCCTCTCTATGGAACCTATGGGAAACGCCGCTCCA 960
 961 CAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACA 1000
 1001 GAACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATAT 1040

5 1041 CGGTATCAACAACCAGCAACTTCCGTTCTTGACGGAACA 1080

1081 GAGTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTG 1120

10 1121 TTTACAGAAAGAGCGGAACCGTTGATTCCTTGGACGAAAT 1160

1161 CCCACCACAGAACAACAATGTGCCACCCAGGCAAGGATTC 1200

15 1201 TCCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGAT 1240

1241 TCAGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTT 1280

1281 CTCTTGGATACACCGTAGTGCTGAGTTCAACAACATCATC 1320

25 1321 GCATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAA 1360

1361 ACTTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATT 1400

30 1401 CACTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAAT 1440

1441 AACATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACT 1480

35 1481 TCCCATCCACATCTACCAGATATAGAGTTCTGTGTGAGGTA 1520

40 1521 TGCTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGT 1560

1561 AATTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTA 1600

45 1601 CCTCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTT 1640

50

55

1641 TGAAAGTGCCAATGCTTTTACATCTTCACTCGGTAAACATC 1680
 1681 GTGGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTA 1720
 1721 TCGACAGATTTCGAGTTCATTCCAGTTACTGCAACACTCGA 1760
 1761 GGCTGAG 1767.

C. Un gène de structure codant une protéine insecticide de *B.t.k.* HD-1 comportant la séquence :

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 161 CTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200
 201 CTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT 240
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360

5 361 CCTACTAACCCAGCTCTCCGCCGAGGAAATGCGTATTCAAT 400
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 10 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCTG 480
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
 15 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTCCGATGCTGC 560
 20 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
 601 GGAAACTACACCGACCACGCTGTTGTTGGTACAACACTG 640
 25 641 GCTTGSAGCGTGCTGCGGGTCTGATTCTAGAGATTGGAT 680
 30 681 TAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT 720
 721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760
 35 761 CCTACCCCTATCCGTACAGTGTCCTCAACTTACCAGAGAAAT 800
 801 CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840
 40 841 CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880
 45 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920
 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960

50

55

961 ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTA 1000
 1001 CCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCACA 1040
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080
 1081 ACCTTGTCCTTCCACCTTGTACAGAAGACCCTTCAATATCG 1120
 1121 GATCAACAACAGCAACTTTCCGTTCTTGACGGAACAGA 1160
 1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200
 1201 TACAGAAAGAGCGGAACCGTTGATTCCTTGGACGAAATCC 1240
 1241 CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC 1280
 1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC 1320
 1321 AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT 1360
 1361 CATGGATTCATCGTAGTGCTGAGTTCAACAATATCATTCC 1400
 1401 TTCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT 1440
 1441 AACCTTGGATCTGGAACCTTCTGTGCTGAAAGGACCAGGCT 1480
 1481 TCACAGGAGGTGATATTCTTAGAAGAACTTCTCCTGGCCA 1520
 1521 GATTAGCACCCCTCAGAGTTAACATCACTGCACCACTTTCT 1560

1561 CAAAGATATCGTGTCTAGGATTCTGTTACGCATCTACCACTA 1600
 1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640
 1641 TCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680
 1681 TTGCAATCCGGCAGCTTCAGAACCCTCGGTTTCACTACTC 1720
 1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTACCCCTTAG 1760
 1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC 1800
 1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCTTCGAGGCTG 1840
 1841 AGTAC 1845.

D. Un gène de structure codant une protéine insecticide dérivée de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40
 41 ACTGCTTGAGTAAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 161 CTGGGTTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200

5 201 CTTTGGTCCATCTCAATGGGATGCATTCCCTGGTGCAAATT 240
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCCGCCAGGA 280
 10 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360
 15 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400
 20 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG 480
 25 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560
 30 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
 35 601 GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTG 640
 641 GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT 680
 40 681 TAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT 720
 721 TTGGACATTGTGTCTCTCTCCCGAACTATGACTCCAGAA 760
 45 761 CCTACCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800

50

55

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
	841	CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
10	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
15			
	961	ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTA	1000
20			
	1001	CCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCACA	1040
25	1041	ACAACGTATCGTTGCTCAACTAGGTGAGGGTGTCTACAGA	1080
	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG	1120
30			
	1121	GTATCAACAACCAGCAACTTCCGTTCTTGACGGAACAGA	1160
35	1161	GTTGCGCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
	1201	TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCC	1240
40	1241	CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC	1280
	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC	1320
45			
	1321	AGCAACAGTTCCGTTGAGCATCATCAGAGCTCCTATGTTCT	1360
50	1361	CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC	1400

55

1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC 1440
 1441 TTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA 1480
 1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520
 1521 CATTCAGAAATAGAGGGTATATTGAAGTTCCAATTCACCTC 1560
 1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600
 1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640
 1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680
 1681 TCCTTGGATAATCTCCAATCCAGCGATTTGCGTTACTTTG 1720
 1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT 1760
 1761 GGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800
 1801 GACAGATTCGAGTTCATTCCAGTTACTGCAACACTCGAGG 1840
 1841 CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTAATGCG 1880
 1881 CTGTTTACGTCTACAAACCAGCTTGGACTCAAGACAAATG 1920
 1921 G 1921;

E. Un gène de structure codant la protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40
 41 ACTGCTTGAGTAACCCAGAGTTGAAGTACTTGGTGGAGA 80
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 161 CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200
 201 CTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT 240
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360
 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG 480
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTGATGCTGC 560
 561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600

5
 601 GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTG 640
 641 GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT 680
 10
 681 TAGATACAACCAAGTTCAGGAGAGAATTGACCCTCACAGTT 720
 721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760
 15
 761 CCTACCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800
 801 CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840
 841 CGTGGTTCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880
 25
 881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920
 921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960
 30
 961 ATCATGGCCTCTCCAGTTGGATTTCAGCGGGCCCGAGTTTA 1000
 35
 1001 CCTTTCCTCTCTATGGAACATATGGGAAACGCCGCTCCACA 1040
 1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080
 40
 1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG 1120
 45
 1121 GSTATCAACAACCAGCAACTTTCGGTTCTTGACGGAACAGA 1160
 1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200

50

55

1201 TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCC 1240
1241 CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC 1280
1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC 1320
1321 AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT 1360
1361 CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC 1400
1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC 1440
1441 TTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA 1480
1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520
1521 CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCAC TTC 1560
1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600
1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640
1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680
1681 TCCTTGGATAATCTCCAATCCAGCGATTTCCGTTACTTTG 1720
1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT 1760
1761 GGGTGT TAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800

5
 1801 GACAGATTCCGAGTTCATTCCAGTTACTGCAACACTCGAGG 1840
 1841 CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC 1880
 10
 1881 GCTGTTTACGTCTACAAACCAGCTCGGCCTCAAGACCAAT 1920
 1921 GTGACGGATTATCATATTGATCAAGTGTCCAACCTTGGTGA 1960
 15
 1961 CCTACCTCAGCGATGAGTTCTGTCTGGATGAAAAGCGAGA 2000
 20
 2001 ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT 2040
 2041 GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA 2080
 25
 2081 ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT 2120
 30
 2121 TACCATCCAGGGAGGTGACGACGTGTTCAAGGAGAACTAC 2160
 2161 GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT 2200
 35
 2201 ACCTCTACCAGAAGATCGACGAGTCCAAGTTGAAAGCCTT 2240
 40
 2241 TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA 2280
 2281 GACCTCGAGATCTACCTCATCCGCTACAATGCAAAACATG 2320
 45
 2321 AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT 2360
 50
 2361 TTCAGCCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT 2400

55

5
 2401 CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT 2440
 2441 GTTCGTGTAGGGATGGAGAAAAGTGTGCCCCATCATTGCGCA 2480
 10
 2481 TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA 2520
 2521 AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAAGA 2560
 15
 2561 CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT 2600
 2601 CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG 2640
 2641 AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAGAAGT 2680
 25
 2681 TGGATGGGAGACCAACATCGTCTACAAAGAGGCAAAGA 2720
 2721 ATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAA 2760
 30
 2761 TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG 2800
 35
 2801 ATAAACGTGTTTCATAGCATTGCGAGAAGCTTATCTGCCTGA 2840
 2841 GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTTTGAA 2880
 40
 2881 GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTCTACG 2920
 45
 2921 ATGCCAGAAACGTCATCAAGAACGGTGACTTCAACAATGG 2960
 2961 CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA 3000

50

55

5 3001 GAACAAAACAACCAACGTTCCGGTCCTTGTTGTTCCGGAAT 3040
 3041 GGGGAAGCAGAAGTGTACACAAGAAGTTCGTGTCTGTCCGGG 3080
 10 3081 TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGA 3120
 3121 TATGGAGAAGGTTGCGTAACCATTTCATGAGATCGAGAACA 3160
 15 3161 ATACAGACGAAGTGAAGTTTAGCAACTGCGTAGAAGAGGA 3200
 20 3201 AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT 3240
 3241 GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA 3280
 25 3281 ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATTA 3320
 3321 TGCCTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA 3360
 30 3361 AGAGAGAATCCTTGTGAATTTAACAGAGGGTATAGGGATT 3400
 35 3401 ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA 3440
 3441 ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA 3480
 40 3481 GAAACGGAAGGAACATTTATCGTGGACAGCGTGAATTAC 3520
 45 3521 TCCTTATGGAGGAA 3534.

50 F. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

55

1 ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA 40
 5
 41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
 10
 81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
 121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
 15
 161 CTGGGTTCTGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200
 201 CTTTGGTCCATCTCAATGGGATGCATTCTTGGTGCAAATT 240
 20
 241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280
 25
 281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
 321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGAAGCCGAT 360
 30
 361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400
 401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
 35
 441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG 480
 40
 481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
 521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTTCGATGCTGC 560
 45
 50
 55

5

561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600

601 GGAAACTACACCGACCACGCTGTTCGTTGGTACAACACTG 640

10

641 GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT 680

681 TAGATAACAACAGTTCAGGAGAGAATTGACCCTCACAGTT 720

15

721 TTGGACATTGTGTCTCTCTCCCGAACTATGACTCCAGAA 760

761 CCTACCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800

20

801 CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC 840

841 CGTGGTTCCTGCCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880

25

881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920

921 CGATGCTCACAGAGGAGATATTACTGGTCTGGACACCAG 960

30

961 ATCATGGCCTCTCCAGTTGGATTGACGGGCCCCGAGTTTA 1000

35

1001 CCTTTCCTCTCTATGGAACCTATGGGAAACGCCGCTCCACA 1040

1041 ACAACGSTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080

40

1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG 1120

45

1121 GTATCAACAACCAGCAACTTCCGTTCTTGACGGAACAGA 1160

1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200

50

1201 TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCC 1240

1241 CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC 1280

55

1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC 1320

5	1321	AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCT	1360
	1361	CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC	1400
10	1401	ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	1441	TTTCTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCA	1480
15	1481	CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA	1520
	1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCATTCT	1560
20	1561	CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG	1600
	1601	CTTCTGTGACCCCTATTCACTCAACGTTAATTGGGGTAA	1640
25	1641	TTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC	1680
	1681	TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG	1720
30	1721	AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT	1760
35	1761	GGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC	1800
	1801	GACAGATTCGAGTTCATTCCAGTTACTGCAACACTCGAGG	1840
40	1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
	1881	GCTGTTTACGTCTACAAACCACTAGGGCTAAAAACAAAT	1920
45	1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA	1960
	1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
50	2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
55	2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080

5

2081 ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT 2120

2121 TACCATCCAAGGAGGGGATGACGTATTTAAAGAAAATTAC 2160

10

2151 GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT 2200

2201 ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT 2240

15

2241 TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA 2280

2281 GACTTAGAAATCTATTTAATTGCTACAATGCAAAACATG 2320

20

2321 AAACAGTAAATGTGCCAGGTACGGGTTTCCTTATGGCCGCT 2360

25

2361 TTCAGCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT 2400

2401 CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT 2440

30

2441 GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATCATTGCGCA 2480

2481 TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA 2520

35

2521 AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA 2560

2561 CGCAAGATGGGCACGCCAAGACTAGGGAATCTAGAGTTTCT 2600

40

2601 CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG 2640

45

2641 AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT 2680

2681 TGGAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA 2720

50

2721 ATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAA 2760

2761 TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG 2800

55

2801 ATAAACGTGTTTCATAGCATTGAGAGAAGCTTATCTGCCTGA 2840

2841 GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTTGAA 2880
 2881 GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG 2920
 2921 ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG 2960
 2961 CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA 3000
 3001 GAACAAAACAACCAACGTTTCGGTCCTTGTTGTTCCGGAAT 3040
 3041 GGAAGCAGAAGTGTCAAGAAGTTCGTGCTGTCCGGG 3080
 3081 TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGA 3120
 3121 TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA 3160
 3161 ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA 3200
 3201 AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT 3240
 3241 GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA 3280
 3281 ATCGAGGATATAACGAAGCTCCTTCGGTACCAGCTGATTA 3320
 3321 TCGTTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA 3360
 3361 AGAGAGAATCCTTGTTGAATTTAACAGAGGGTATAGGGATT 3400
 3401 ACACGCCACTACCAAGTTGGTTATGTGACAAAAGAATTAGA 3440
 3441 ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA 3480
 3481 GAAACGGAAGGAACATTTATCGTGGACAGCGTGGAATTAC 3520
 3521 TCCTTATGGAGGAA 3534.

G. Un gène de structure codant une protéine insecticide en pleine longueur de *B.t.k.* HD-73 comportant la séquence :

```

5      1  ATGGACAACAACCCAAACATCAACGAATGCATTCCATACA  40
      .
      .
10     41  ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA  80
      .
      .
      81  ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG  120
      .
      .
15     121  TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG  160
      .
      .
      161  CTGGGTTCGTTCTCGGACTAGTTGACATCATCTGGGGTAT  200
      .
      .
20     201  CTTTGGTCCATCTCAATGGGATGCATTCCCTGGTGCAAATT  240
      .
      .
      241  GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA  280
      .
      .
25     281  ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA  320
      .
      .
      321  CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT  360
      .
      .
30     361  CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT  400
      .
      .
35     401  TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT  440
      .
      .
      441  GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG  480
      .
      .
      481  TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG  520
      .
      .
45     521  ACGTTAGCGTGTTTTGGGCAAAGGTGGGGATTGATGCTGC  560
      .
      .
      561  AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT  600
      .
      .
50     601  GGAAACTACACCGACACGCTGTTGTTGGTACAACACTG  640
      .
      .
      641  GCTTGGAGCGTGTCTGGGGTCCTGATTCTAGAGATTGGAT  680

```

5

681 TAGATACAACCCAGTTCAGGAGAGAATTGACCCCTCACAGTT 720

721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760

10

761 CCTACCCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800

801 CTATACTAACCCAGTTCCTTGAGAACTTCGACGGTAGCTTC 840

15

841 CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC 880

881 CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC 920

20

921 CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG 960

961 ATCATGGCCTCTCCAGTTGGATTTCAGCGGGCCCGAGTTTA 1000

25

1001 CCTTTCCTCTCTATGGAACCTATGGGAAACGCCGCTCCACA 1040

30

1041 ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA 1080

1081 ACCTTGTCTTCCACCTTGTACAGAAGACCCCTTCAATATCG 1120

35

1121 GTATCAACAACCAGCAACTTTCCGTTCTTGACGGAACAGA 1160

1161 GTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT 1200

40

1201 TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCC 1240

1241 CACCACAGAACACAATGTGCCACCCAGGCAAGGATTCTC 1280

45

1281 CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC 1320

50

1321 AGCAACAGTTCGGTGAGCATCATCAGAGCTCCTATGTTCT 1360

1361 CTTGGATACACCGTAGTGCTGAGTTCAACAACATCATCGC 1400

55

1401 ATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC 1440

5

1441 TTTCTCTTCAACGGTCTGTCAATTCAGGACCAGGATTCA 1480

1481 CTGGTGGAGACCTCGTTAGACTCAACAGCAGTGGAAATAA 1520

10

1521 CATTTCAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC 1560

1561 CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATG 1600

15

1601 CTTCTGTGACCCCTATTACCTCAACGTTAATTGGGGTAA 1640

1641 TTCATCCATCTTCTCCAATACAGTTCAGCTACAGCTACC 1680

20

1681 TCCTTGGATAATCTCCAATCCAGCGATTTCGGTTACTTTG 1720

1721 AAAGTGCCAATGCTTTTACATCTTCACTCGGTAACATCGT 1760

25

1761 GGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC 1800

1801 GACAGATTTCGAGTTCATTCCAGTTACTGCAACACTCGAGG 1840

30

1841 CTGAGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGC 1880

1881 CCTCTTTACCTCCACCAATCAGCTTGGCTTGAAAACATAAC 1920

35

1921 GTTACTGACTATCACATTGACCAAGTGTCCAACCTTGGTCA 1960

1961 CCTACCTTAGCGATGAGTCTGCCTCGACGAGAAGCGTGA 2000

40

2001 ACTCTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGAC 2040

45

2041 GAGAGGAATCTCTTGCAAGACTCCAACCTTCAAAGACATCA 2080

2081 ACAGGCAGCCAGAACGTGGTTGGGGTGAAGCACCGGGAT 2120

50

2121 CACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC 2160

55

2161 GTCACCCCTCTCCGGAACCTTTCGACGAGTGCTACCCTACCT 2200

5

2201 ACTTGTACCAGAGATCGATGAGTCCAAACTCAAAGCCTT 2240

10

2241 CACCAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA 2280

2281 GACCTTGAAATCTACTCGATCAGGTACAATGCCAAGCACG 2320

2321 AGACCGTGAATGTCCCAGGTACTGGTTCCTCTGGCCACT 2360

15

2361 TTCTGCCCAATCTCCCATGGGAAGTGTGGAGAGCCTAAC 2400

2401 AGATGCGCTCCACACCTTGAGTGGAAATCCTGACTTGGACT 2440

20

2441 GCTCCTGCAGGGATGGCGAGAAGTGTGCCCACCATTCTCA 2480

2481 TCACTTCTCCTTGGACATCGATGTGGGATGTACTGACCTG 2520

25

2521 AATGAGGACCTCGGAGTCTGGGTCACTTCAAGATCAAGA 2560

30

2561 CCCAAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCT 2600

2601 CGAAGAGAAACCATTGGTCGGTGAAGCTCTCGCTCGTGTG 2640

35

2641 AAGAGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAAAC 2680

2681 TCGAATGGGAAACTAACATCGTTTACAAGGAGGCCAAAGA 2720

40

2721 GTCCGTGGATGCTTTGTTGTTGTAAGTCCCAATATGATCAG 2760

2761 TTGCAAGCCGACACCAACATCGCCATGATCCACGCCGCAG 2800

45

2801 ACAAACGTGTGCACAGCATTCTGTGAGGCTTACTTGCCTGA 2840

50

2841 GTTGTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAG 2880

2881 GAACTTGAGGGACGTATCTTTACCGCATTCTCCTTGTACG 2920

55

2921 ATGCCAGAAACGTCATCAAGAACGGTGACTTCAACBAATGG 2960

2961 CCTCAGCTGCTGGAATGTGAAAGGTCATGTGGACGTGGAG 3000
 3001 GAACAGAACAAATCAGCGTTCCGTCCTGGTTGTGCCTGAGT 3040
 3041 GGGGAAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGG 3080
 3081 TAGAGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGA 3120
 3121 TACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACA 3160
 3161 ACACCGACGAGCTTAAGTTCTCCAACTGCGTCGAGGAAGA 3200
 3201 AATCTATCCCAACAACACCGTTACTTGCAACGACTACACT 3240
 3241 GTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTA 3280
 3281 ACAGAGGTTACAACGAAGCTCCTTCCGTTCCCTGCTGACTA 3320
 3321 TGCCTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGA 3360
 3361 CGTGAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACT 3400
 3401 ACACACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGA 3440
 3441 GTACTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGT 3480
 3481 GAAACCGAGGGGAACCTTCATCGTGGACAGCGTGGAGCTTC 3520
 3521 TCTTGATGGAGGAA 3534.

H. Un gène de structure qui code une protéine insecticide de *B.t.t.* Comportant La séquence :

5

1 ATGACTGCAGACAACAACACCGAAGCCCTCGACAGTTCTA 40

41 CCACTAAGGATGTTATCCAGAAGGGTATCTCCGTTGTGGG 80

10

81 AGACCTCTTGGGCGTGGTTGGATTTCCTTCGGTGGAGCC 120

121 CTCGTGAGCTTCTATACAACTTTCTCAACACCATTTGGC 160

15

161 CAAGCGAGGACCCCTTGGAAAGCATTTCATGGAGCAAGTTGA 200

20

201 AGCTCTTATGGATCAGAAGATTGCAGATTATGCCAAGAAC 240

241 AAGGCTTTGGCAGAACTCCAGGGCCTTCAGAACAAATGTGG 280

25

281 AGGACTACGTGAGTGCATTGTCCAGCTGGCAGAAGAACCC 320

30

321 TGTTAGCTCCAGAAATCCTCACAGCCAAGGTAGGATCAGA 360

361 GAGTTGTTCTCTCAAGCCGAATCCCACTTCAGAAATTCCA 400

35

401 TGCCTAGCTTTGCTATCTCCGGTTACGAGGTTCTTTTCCT 440

441 CACTACCTATGCTCAAGCTGCCAACACCCACTTGTTTCTC 480

40

481 CTTAAGGACGCTCAAATCTATGGAGAAGAGTGGGGATACG 520

45

521 AGAAAGAGGACATTGCTGAGTTCTACAAGCGTCAACTTAA 560

561 GCTCACC CAAGAGTACACTGACCATTGCGTGAAATGGTAT 600

50

601 AACGTTGGTCTCGATAAGCTCAGAGGCTCTTCCTACGAGT 640

641 CTTGGGTGAACCTTCAACAGATACAGGAGAGAGATGACCTT 680

55

5
681 GACTGTGCTCGATCTTATCGGCACTCTTTCCTTGTACGAT 720

721 GTGAGACTCTACCCAAAGGAAGTGAAAAGTGAGCTTACCA 760

10
761 GAGACGTGCTCACTGACCCTATTGTGGGAGTCAACAACCT 800

801 TAGGGGTTATGGAACTACCTTCAGCAATATCGAAAAGTAC 840

15
841 ATTAGGAAACCACATCTCTTCGACTATCTTCACAGAATTC 880

881 AATTCCACACAAGGTTTCACACCAGGATACATGGTAACGA 920

20
921 CTCCTTCAACTATTGGTCCGGTAAGTATGTTTCCACCAGA 960

961 CCAAGCATTTGGATCTAATGACATCATCACATCTCCCTTCT 1000

25
1001 ATGGTAACAAGTCCAGTGAACCTGTGCAGAACCTTGAGTT 1040

30
1041 CAACGGCGAGAAAAGTCTATAGAGCCGTGCAAAACACCAAT 1080

1081 CTCGCTGTGTGGCCATCCGCAGTTTACTCAGGCGTCACAA 1120

35
1121 AGGTGGAGTTTATGTCAGTATAACGATCAGACCGATGAGGC 1160

1161 CAGCACCAGACTTACGACTCCAAACGTAACGTTGGCGCA 1200

40
1201 GTCTCTTGGGATTCTATCGACCAATTGCCTCCAGAAACCA 1240

1241 CAGACGAACCATTTGAGAGGGGCTACAGCCACCAACTTAA 1280

45
1281 CTATGTGATGTGCTTCTTGATGCAAGGTTCCAGAGGGACC 1320

1321 ATTCCAGTGTTGACCTGGACACACAAGTCCGTGGACTTCT 1360

50
1361 TCAACATGATCGATAGCAAGAAGATCACTCAACTTCCCTT 1400

55
1401 GGTGAAAGCCTACAAGCTGCAATCTGGTGCTTCCGTTGTC 1440

1441 GCAGGTCCCAGATTCACTGGAGGTGACATCATCCAGTGCA 1480
 1481 CAGAGAACGGCAGCGCAGCTACTATCTACGTGACACCTGA 1520
 1521 TGTGTCTTACTCTCAGAAGTACAGGGCACGTATTCATTAC 1560
 1561 GCATCTACCAGCCAGATCACCTTCACACTCAGCTTGGATG 1600
 1601 GAGCACCCCTTCAACCAGTATTACTTTGACAAGACCATCAA 1640
 1641 CAAAGGTGACACTCTCACATACAATAGCTTCAACTTGGCA 1680
 1681 AGTTTCAGCACACCATTGAACTCTCAGGCAACAATCTTC 1720
 1721 AGATCGGCGTCACCGGTCTCAGCGCCGGAGACAAAGTCTA 1760
 1761 CATCGACAAGATTGAGTTTCATCCCAGTGAAC 1791

I. Un gène de structure qui code une protéine insecticide de *B.t. entomocidus* comportant la séquence :

1 ATGGAGGAGAACAACCAAAACCAATGCATTCCATACAACT 40
5
41 GCTTGAGTAACCCAGAAGAGGTATTGCTTGATGGAGAACG 80
81 CATTTC AACCGGTAACCTCTCCATCGACATCTCCTTGTCC 120
10
121 TTGGTCCAGTTTCTGGTCAGCAACTTCGTGCCAGGTGGTG 160
15
161 GGTTCCCTTGTCGGACTAATTGACTTCGTCTGGGGTATCGT 200
201 TGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATTGAG 240
20
241 CAGTTGATCAACGAGAGGATCGCTGAGTTCGCCAGGAACG 280
25
281 CTGCCATCGCTAACTTGGAAGGATTGGGCAATAACTTCAA 320
321 CATCTATGTGGAGGCCTTCAAAGAGTGGGAAGAGGACCCT 360
30
361 AACCAACCCAGAGACCCGCACTAGGGTGATCGACAGATTCA 400
401 GAATCTTGACGGCCTCTTGGAGAGAGATATCCCATCCTT 440
35
441 CAGAATCTCTGGCTTCGAAGTTCCTCTCTTGTCCGTGTAC 480
40
45
50
55

5

481 GCTCAAGCAGCTAATCTTCACCTCGCTATCCTTCGAGACA 520

521 GTGTCATCTTTGGGGAAAGGTGGGGATTGACCACTATCAA 560

10

561 CGTCAATGAGAATTACAACAGACTTATCAGGCACATTGAC 600

601 GAGTACGCCGACCACTGTGCTAACACCTACAACCGTGGCT 640

15

641 TGAACAATCTCCCTAAGTCTACTTATCAAGATTGGATTAC 680

681 CTACAACAGGTTGAGGAGAGACTTGACCCTCACAGTTTTG 720

20

721 GACATTGCAGCTTTCTTCCGAACTATGACAACAGGAGAT 760

761 ACCCTATCCAACCACTGGGTCAACTTACCAGAGAAGTCTA 800

25

801 TACTGACCCACTTATCAACTTCAACCCTCAGTTGCAAAGT 840

30

841 GTCGCCCAACTTCCCACATTCAACGTCATGGAGTCCAGCC 880

881 GTATCAGGAACCCACACTTGTTTGACATCTTGAACAACCT 920

35

921 TACTATCTTCACCGATTGGTTCAGCGTTGGGCGTAACTTC 960

961 TATTGGGGTGGACACAGGGTCATCTCCTCTCTTATTGGAG 1000

40

1001 GTGGGAACATTACCTCTCCTATCTATGGACGTGAGGCAAA 1040

1041 CCAGGAGCCACCACGTAGTTTCACCTTCAACGGTCCAGTC 1080

45

1081 TTCAGAACCTTGTCTAACCCCTACCTTGAGATTGCTCCAGC 1120

50

1121 AACCTTGGCCAGCTCCACCTTTCAACCTTAGAGGTGTTGA 1160

1161 GGGCGTTGAGTTCTCTACTCCTACCAACTCCTTCACTTAC 1200

55

1201 AGAGGTAGAGGAACCGTTGATTCTTGACCGAACTCCAC 1240

5

1241 CAGAGGACAATAGCGTGCCACCCAGGGAAGGCTACTCCCA 1280

1281 CAGGTTGTGCCACGCAACCTTCGTGCAGCGTTCCGGAACT 1320

10

1321 CCATTCCTCACTACAGGAGTTGTGTTCTCATGGACTGATC 1360

1361 GTAGTGCTACTCTCACTAATACCATTGATCCCGAGAGGAT 1400

15

1401 CAATCAAATCCCATTTGGTCAAGGGTTTCCGTGTGTGGGGA 1440

1441 GGAACCTTCTGTATCACAGGACCAGGCTTCACAGGAGGTG 1480

20

1481 ATATTCTTAGAAGAAACACTTTTGGCGACTTTGTGAGCCT 1520

25

1521 CCAAGTTAACATCAACTCTCCAATTACTCAAAGATATCGT 1560

1561 CTCAGGTTTCGTTACGCATCTTCCCGTGACGCTAGAGTCA 1600

30

1601 TCGTGCTCACCGGAGCAGCTTCTACCGGTGTGCGTGGACA 1640

1641 AGTCTCOGTGAACATGCCACTCCAGAAGACTATGGAGATC 1680

35

1681 GGCGAGAACTTGACATCCAGGACCTTCAGATACACCGACT 1720

1721 TCTCTAACCCTTTCAGTTTCCGTGCCAACCCCTGACATCAT 1760

40

1761 TGSCATTAGCGAACAACCTCTCTTTGGAGCTGGTAGCATC 1800

45

1801 TCATCTGGCGAATTGTACATTGACAAGATTGAGATCATTC 1840

1841 TTGCCGACGCTACCTTCGAGGCTGAGTCTGACCTTGAGAG 1880

50

1881 AGCCCAGAAGGCTGTGAACGCCCTCTTTACCTCCTCTAAT 1920

1921 CAGATTGGCTTGAAAACCTGACGTTACTGACTATCACATTG 1960

55

1961 ACCAAGTGTCCAACCTTGGTGGACTGCCCTTAGCGATGAGTT 2000

5

2001 CTGCCTCGACGAGAAGCGTGAACTCTCCGAGAAAGTTAAA 2040

2041 CACGCCAAGCGTCTCAGCGACGAGAGGAATCTCTTGCAAG 2080

10

2081 ACCCCAACTTCAGAGGCATCAACAGGCAGCCAGACCGTGG 2120

2121 TTGGAGAGGAAGCACCGACATCACCATCCAAGGAGGCGAC 2160

15

2161 GATGTGTTCAAGGAGAACTACGTCACCCCTCCCAGGAACTG 2200

2201 TGGACGAGTGCTACCCCTACCTACTTGTACCAGAAGATCGA 2240

20

2241 TGAGTCCAAACTCAAAGCCTACACCAGGTATGAACTTAGA 2280

2281 GGCTACATCGAAGACAGCCAAGACCTTGAAATCTACCTCA 2320

25

2321 TCAGGTACAATGCCAAGCAGAGATCGTGAATGTCCCAGG 2360

30

2361 TACTGGTTCCCTCTGGCCACTTTCTGCCCAAATGCCCAT 2400

2401 GGGAAAGTGTGGAGAGCCTAACAGATGCGCTCCACACCTTG 2440

35

2441 AGTGGAAATCCTGACTTGGACTGCTCCTGCAGGGATGGCGA 2480

2481 GAAGTGTGCCCACCATTCATCACTTCACCTTGGACATC 2520

40

2521 GATGTGGGATGTACTGACCTGAATGAGGACCTCGGAGTCT 2560

2561 GGGTCATCTTCAAGATCAAGACCCAAGACGGACACGCAAG 2600

45

2601 ACTTGGCAACCTTGAGTTTCTCGAAGAGAAACCATTGCTC 2640

50

2641 GGTGAAGCTCTCGCTCGTGTGAAGAGAGCAGAGAAGAAGT 2680

2681 GGAGGGACAAACGTGAGAACTCCAACTCGAGACTAACAT 2720

55

2721 CGTTTACAAGGAGGCCAAAGAGTCCGTGGATGCTTTGTTC 2760

5

2761 GTGAACTCCCAATATGATAGGTTGCAAGTGGACACCAACA 2800

2801 TCGCCATGATCCACGCTGCAGACAAACGTGTGCACAGGAT 2840

10

2841 TCGTGAGGCTTACTTGCCTGAGTTGTCCGTGATCCCTGGT 2880

2881 GTGAACGCTGCCATCTTCGAGGAACTTGAGGGACGTATCT 2920

15

2921 TTACCGCATACTCCTTGTACGATGCCAGAAACGTCATCAA 2960

2961 GAACGGTGACTTCAACAATGGCCTCTTGTGCTGGAATGTG 3000

20

3001 AAAGGTCATGTGGACGTGGAGGAACAGAACAAATCACCGTT 3040

3041 CCGTCCTGGTTATCCCTGAGTGGGAAGCTGAAGTGTCCCA 3080

25

3081 AGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGT 3120

30

3121 GTGACCGCTTACAAGGAGGGATACGGTGAAGGTTGCGTGA 3160

3161 CCATCCACGAGATCGAGGACAACACCGACGAGCTTAAGTT 3200

35

3201 CTCCAAGTGCCTCGAGGAAGAAGTCTATCCCAACAACACC 3240

3241 GTTACTTGCAACAACACTACACTGGGACCCAGGAAGAGTACG 3280

40

3281 AAGGTACCTACACTAGCCGTAACCAAGGTTACGACGAAGC 3320

45

3321 TTACGGAAACAATCCTTCCGTTCCCTGCTGACTATGCCTCC 3360

3361 GTGTACGAGGAGAAATCCTACACAGATGGCAGACGTGAGA 3400

50

3401 ACCCTTGCGAGTCCAACAGAGGTTACGGTGAAGTACACACC 3440

3441 ACTTCCAGCAGGCTATGTTACCAAGGACCTTGAGTACTTT 3480

55

3481 CCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAAACCG 3520

3521 AGGGAACCTTCATCGTGGACAGCGTGGAGCTTCCTTGAT 3560

5

3561 GGAGGAA 3567.
|

10

J. Un gène de structure qui code - une protéine insecticide P2 comportant la séquence :

15

20

25

30

35

40

45

50

55

5

1 ATGGACAACAACGTCTTGAACTCTGGTAGAACACCATCT 40

10

41 GCGACGCATACAACGTCGTGGCTCAGGATCCATTCAGCTT 80

15

81 CGAACACAAGAGCCTCGACACTATTCAGAAGGAGTGGATG 120

121 GAATGGAAACGTA CTGACCACTCTCTCTACGTGCGACCTG 160

20

161 TGGTTGGAACAGTGTCCAGCTTCCTTCTCAAGAAGGTCGG 200

201 CTCTCTCATCGGAAAACGTATCTTGTCCGAACTCTGGGGT 240

25

241 ATCATCTTTCCATCTGGGTCCACTAATCTCATGCAAGACA 280

281 TCTTGAGGGAGACCGAACAGTTTCTCAACCAGCGTCTCAA 320

30

321 CACTGATACCTTGGCTAGAGTCAACGCTGAGTTGATCGGT 360

361 CTCCAAGCAAACATTCTGTGAGTTCAACCAGCAAGTGGACA 400

35

401 ACTTCTTGAATCCAACTCAGAATCCTGTGCCTCTTTCCAT 440

441 CACTTCTTCCGTGAACACTATGCAGCAACTCTTCCTCAAC 480

40

481 AGATTGCCTCAGTTTCAGATTCAAGGCTACCAGTTGCTCC 520

45

521 TTCTTCCACTCTTTGCTCAGGCTGCCAACATGCACTTGTC 560

50

561 CTTCAACGTGACGTGATCCTCAACGCTGACGAATGGGGA 600

55

5

601 ATCTCTGCAGCCACTCTTAGGACATACAGAGACTACTTGA 640

641 GGAAC TACACTCGTGATTACTCCA ACTATTGCATCAACAC 680

10

681 TTATCAGACTGCCTTTTCGTGGACTCAATACTAGGCTTCAC 720

721 GACATGCTTGAGTTCAGGACCTACATGTTCTTAACGTGT 760

15

761 TTGAGTACGTCAGCATTTCGGAGTCTCTTCAAGTACCAGAG 800

801 CTTGATGGTGTCTCTGGAGCCAATCTCTACGCCTCTGGC 840

20

841 AGTGGACCACAGCAA ACTCAGAGCTTCACAGCTCAGAACT 880

25

881 GGCCATTCTTG TATAGCTTGTTCCAAGTCAACTCCA ACTA 920

921 CATTCTCAGTGGTATCTCTGGGACCAGACTCTCCATAACC 960

30

961 TTTCCCAACATTGGTGGACTTCCAGGCTCCACTACAACCC 1000

1001 ATAGCCTTAACTCTGCCAGAGTGAACTACAGTGGAGGTGT 1040

35

1041 CAGCTCTGGATTGATTGGTGCAACTAACTTGAACCACAAC 1080

1081 TTCAATTGCTCCACCGTCTTGCCACCTCTGAGCACACCGT 1120

40

1121 TTGTGAGGTCTTGCTTGACAGCGGTACTGATCGCGAAGG 1160

45

1161 AGTTGCTACCTCTACAAACTGGCAAACCGAGTCCTTCCAA 1200

1201 ACCACTCTTAGCCTTCGGTGTGGAGCTTTCTCTGCACGTG 1240

50

1241 GGAATTCAA ACTACTTTCCAGACTACTTCATTAGGAACAT 1280

1281 CTCTGGTGTTCTCTCGTCATCAGGAATGAAGACCTCACC 1320

55

1321 CGTCCACTTCATTACAACCAGATTAGGAACATCGAGTCTC 1360

1361 CATCCGGTACTCCAGGAGGTGCAAGAGCTTACCTCGTGTC 1400
 1401 TGTCCATAACAGGAAGAACAACATCTACGCTGCCAACGAG 1440
 1441 AATGGCACCATGATTCACCTTGCACCAGAAGATTACACTG 1480
 1481 GATTCACCATCTCTCCAATCCATGCTACCCAAGTGAACAA 1520
 1521 TCAGACACGCACCTTCATCTCCGAAAAGTTCGGAAATCAA 1560
 1561 GGTGACTCCTTGAGGTTTCGAGCAATCCAACACTACCGCTA 1600
 1601 GGTACACTTTGAGAGGCAATGGAAACAGCTACAACCTTTA 1640
 1641 CTTGAGAGTTAGCTCCATTGGTAACTCCACCATCCGTGTT 1680
 1681 ACCATCAACGGACGTGTTTACACAGTCTCTAATGTGAACA 1720
 1721 CTACAACGAACAATGATGGCGTTAACGACAACGGAGCCAG 1760
 1761 ATTCAGCGACATCAACATTGGCAACATCGTGGCCTCTGAC 1800
 1801 AACACTAACGTTACTTTGGACATCAATGTGACCCTCAATT 1840
 1841 CTGGAACCTCCATTTGATCTCATGAACATCATGTTTGTGCC 1880
 1881 AACTAACCTCCCTCCATTGTAC 1902 ou

K. Une séquence de gène de structure codant une protéine de fusion comprenant les acides aminés 610 N-terminaux de *B.t.k.* HD-1 et les acides aminés 567 C-terminaux de *B.t.k.* HD-73, ledit gène comportant la séquence :

1 ATGGGCAACAACCCAAACATCAACGAATGCATTCCATACA 40
5
41 ACTGCTTGAGTAACCCAGAAGTTGAAGTACTTGGTGGAGA 80
10
81 ACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG 120
121 TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTG 160
15
161 CTGGGTTCTGTTCTCGGACTAGTTGACATCATCTGGGGTAT 200
20
201 CTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT 240
241 GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGA 280
25
281 ACCAGGCCATCTCTAGGTTGGAAGGATTGAGCAATCTCTA 320
321 CCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT 360
30
361 CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAAT 400
35
401 TCAACGACATGAACAGCGCCTTGACCACAGCTATCCCATT 440
40
45
50
55

441 GTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTC 480
481 TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAG 520
521 ACGTTAGCGTGTTTGGGCAAAGGTGGGGATTTCGATGCTGC 560
561 AACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT 600
601 GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTG 640
641 GCTTGGAGCGTGTCCTGGGGTCTGATTCTAGAGATTGGAT 680
681 TAGATACAACCAAGTTCAGGAGAGAATTGACCCTCACAGTT 720
721 TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAA 760
761 CCTACCCTATCCGTACAGTGTCCCAACTTACCAGAGAAAT 800

5	801	CTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC	840
	841	CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCC	880
10	881	CACACTTGATGGACATCTTGAACAGCATAACTATCTACAC	920
	921	CGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG	960
15	961	ATCATGGCCTCTCCAGTTGGATTTCAGCGGGCCCGAGTTTA	1000
	1001	CCTTTCCTCTCTATGGAAGTATGGGAAACGCCGCTCCACA	1040
20	1041	ACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA	1080
25	1081	ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCG	1120
	1121	GTATCAACAACCAGCAACTTTCGGTTCTTGACGGAACAGA	1160
30	1161	GTTGCGCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT	1200
	1201	TACAGAAAGAGCGGAACCGTTGATTCCTTGGACGAAATCC	1240
35	1241	CACCACAGAACAACAATGTGCCACCCAGGCAAGGATTCTC	1280
40	1281	CCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTTC	1320
	1321	AGCAACAGTTCCGTTGAGCATCATCAGAGCTCCTATGTTCT	1360
45	1361	CATGGATTTCATCGTAGTGCTGAGTTCAACAATATCATTCC	1400
	1401	TTCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT	1440
50	1441	AACCTTGGATCTGGAACCTCTGTGCGTGAAAGGACCAGGCT	1480
55	1481	TCACAGGAGGTGATATTCTTAGAAGAACTTCTCCTGGCCA	1520

5

1521 GATTAGCACCCTCAGAGTTAACATCACTGCACCACCTTTCT 1560

1561 CAAAGATATCGTGTGTCAGGATTTCGTTACGCATCTACCACTA 1600

10

1601 ACTTGCAATTCCACACCTCCATCGACGGAAGGCCTATCAA 1640

1641 TCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC 1680

15

1681 TTGCAATCCGGCAGCTTCAGAACCCTCGGTTTCACTACTC 1720

1721 CTTTCAACTTCTCTAACGGATCAAGCGTTTTCAACCCTTAG 1760

20

1761 CGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC 1800

25

1801 CGTATTGAGTTTGTGCCTGCCGAAGTTACCCCTCGAGGCTG 1840

1841 AGTACAACCTTGAGAGAGCCCAGAAGGCTGTGAACGCCCT 1880

30

1881 CTTTACCTCCACCAATCAGCTTGGCTTGAAAACCTAACGTT 1920

1921 ACTGACTATCACATTGACCAAGTGTCCAACCTTGGTCACCT 1960

35

1961 ACCTTAGCGATGAGTTCTGCCTCGACGAGAAGCGTGAACT 2000

2001 CTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGACGAG 2040

40

2041 AGGAATCTCTTGCAAGACTCCAACCTTCAAAGACATCAACA 2080

2081 GGCAGCCAGAACGTGGTTGGGGTGAAGCACCGGGATCAC 2120

45

2121 CATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTACGTC 2160

50

2161 ACCCTCTCCGGAACCTTCGACGAGTGCTACCCTACCTACT 2200

2201 TGTACCAGAAGATCGATGAGTCCAAACTCAAAGCCTTCAC 2240

55

2241 CAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAAGAC 2280

5

2281 CTTGAAATCTACTCGATCAGGTACAATGCCAAGCACGAGA 2320

2321 CCGTGAATGTCCCAGGTACTGGTTCCTCTGGCCACTTTC 2360

10

2361 TGCCCAATCTCCCATTTGGGAAGTGTGGAGAGCCTAACAGA 2400

2401 TGCGCTCCACACCTTGAGTGGAAATCCTGACTTGGACTGCT 2440

15

2441 CCTGCAGGGATGGCGAGAAGTGTGCCACCATTCTCATCA 2480

2481 CTTCTCCTTGGACATCGATGTGGGATGTACTGACCTGAAT 2520

20

2521 GAGGACCTCGGAGTCTGGGTCACTTTCAAGATCAAGACCC 2560

2561 AAGACGGACACGCAAGACTTGGCAACCTTGAGTTTCTCGA 2600

25

2601 AGAGAAACCAATTGGTCGGTGAAGCTCTCGCTCGTGTGAAG 2640

30

2641 AGAGCAGAGAAGAAGTGGAGGGACAAACGTGAGAACTCG 2680

2681 AATGGGAAACTAACATCGTTTACAAGGAGGCCAAAGAGTC 2720

35

2721 CGTGGATGCTTTGTTTCTGTAAGTCCCAATATGATCAGTTG 2760

2761 CAAGCCGACACCAACATCGCCATGATCCACGCCGCAGACA 2800

40

2801 AACGTGTGCACAGCATTCGTGAGGCTTACTTGCCTGAGTT 2840

45

2841 GTCCGTGATCCCTGGTGTGAACGCTGCCATCTTCGAGGAA 2880

2881 CTTGAGGGACGTATCTTTACCGCATTCTCCTTGTACGATG 2920

50

2921 CCAGAAACGTCATCAAGAACGGTGACTTCAACAATGGCCT 2960

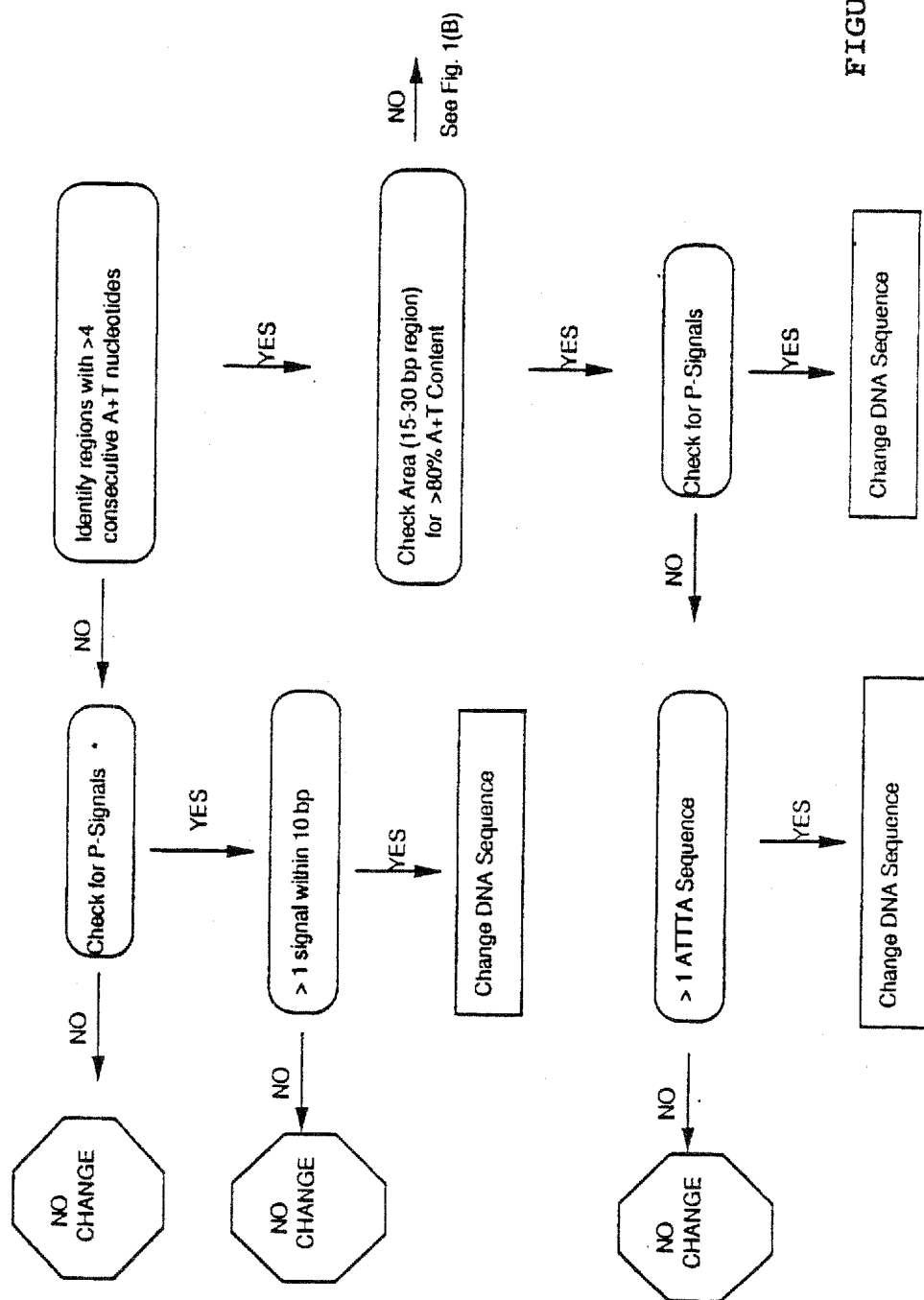
2961 CAGCTGCTGGAATGTGAAAGGTCATGTGGACGTGGAGGAA 3000

55

3001 CAGAACAATCAGCGTTCCGTCCTGGTTGTGCCTGAGTGGG 3040

3041 AAGCTGAAGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTAG 3080
3081 AGGCTACATTCTCCGTGTGACCGCTTACAAGGAGGGATAC 3120
3121 GGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAACA 3160
3161 CCGACGAGCTTAAGTTCTCCAAGTGCCTCGAGGAAGAAAT 3200
3201 CTATCCCAACAACACCGTTACTTGCAACGACTACACTGTG 3240
3241 AATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACA 3280
3281 GAGGTTACAACGAAGCTCCTTCCGTTCCCTGCTGACTATGC 3320
3321 CTCCGTGTACGAGGAGAAATCCTACACAGATGGCAGACGT 3360
3361 GAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACA 3400
3401 CACCACTTCCAGTTGGCTATGTTACCAAGGAGCTTGAGTA 3440
3441 CTTTCCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA 3480
3481 ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTCT 3520
3521 TGATGGAGGAA 3531.

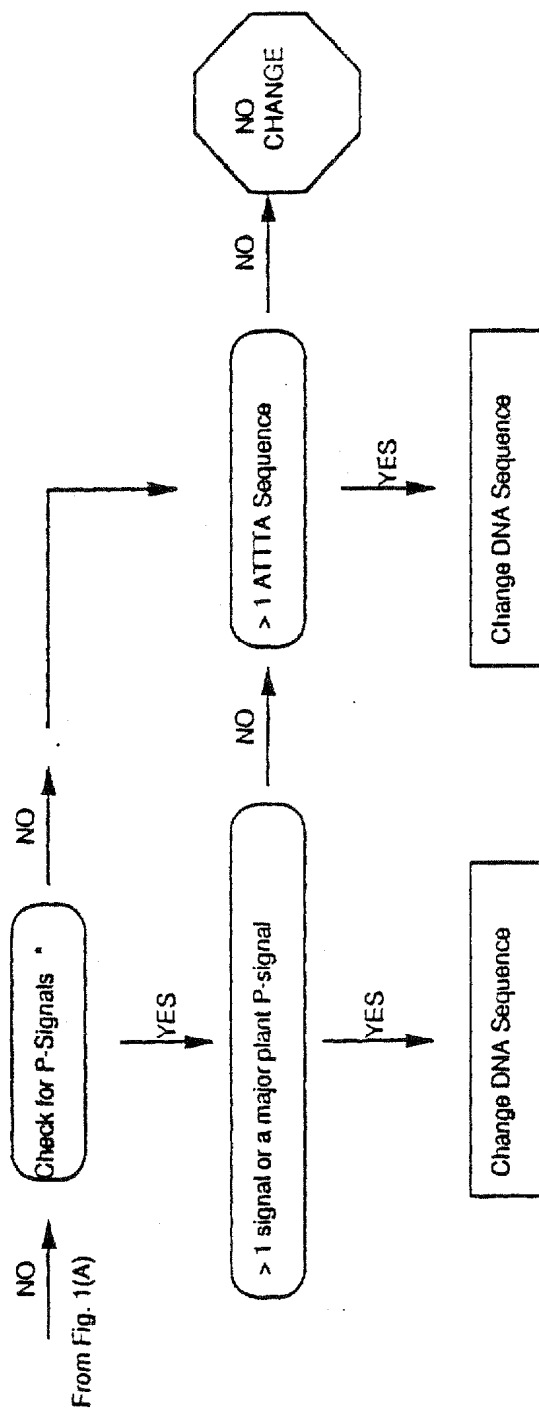
Determination of DNA regions in genes to modify by
site-directed mutagenesis for increased expression in plants



* Polyadenylation Signal Sequences

FIGURE 1 (A)

Determination of DNA regions in genes to modify by
site-directed mutagenesis for increased expression in plants



• Polyadenylation Signal Sequences

FIGURE 1 (B)

1	ATGGCTATAGAACTGGTTACACCCCAATCGATATTTTCCT	40
41	TGTCGCTAACGCAATTTCTTTTGTGAGTGAATTTGTTCCCGG	80
81	TGCTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGA T C	120
121	ATTTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAA	160
161	TTGAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAG C C C G C G	200
201	GAACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTT T	240
241	TATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAG	280
281	ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA	320
321	ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCCT	360
361	CTTTTTGCAGTTCAAATTTATCAAGTTCCTCTTTTATCAG CC C C	400
401	TATATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAG G C C CC C CC C	440
441	AGATGTTTCAGTGTGTTGGACAAAGGTGGGGATTTGATGCC	480
481	GCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTA	520
521	TTGGCAACTATACAGATCATGCTGTACGCTGGTACAATAC	560
561	GGGATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGG	600
601	ATAAGATATAATCAATTTAGAAGAGAATTAACACTAACTG C G C C G C GC T	640
641	TATTAGATATCGTTTCTCTATTTCCGAACCTATGATAGTAG	680
681	AACGTATCCAATTCTGAACAGTTTCCCAATTAACAAGAGAA	720

FIGURE 2A

721	ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT	760
761	TTCGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAG	800
801	TCCACATTTGATGGATATACTTAATAGTATAACCATCTAT	840
841	ACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATC	880
	C C C T C	
881	AAATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATT	920
	G C	
921	CACTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCA	960
961	CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATA	1000
1001	GAACATTATCGTCCACCTTATATAGAAGACCTTTTAATAT	1040
	C	
1041	AGGGATAAATAATCAACAACCTATCTGTTCTTGACGGGACA	1080
	C C C C	
1081	GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG	1120
1121	TATACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAAT	1160
1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT	1200
1201	AGTCATCGATTAAGCCATGTTTCAATGTTTCGTTTCAGGCT	1240
1241	TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT	1280
1281	CTCTTGATACATCGTAGTGCTGAATTTAATAATATAATT	1320
	G C C C C C	
1321	CCTTCATCACAAATTACACAAATACCTTTAACAAAATCTA	1360
	C C C AC C C G	
1361	CTAATCTTGGCTCTGGAACCTTCTGTCGTTAAAGGACCAGG	1400

FIGURE 2B

1401	ATTTACAGGAGGAGATATTCTTCGAAGAACTTCACCTGGC	1440
1441	CAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT	1480
1481	CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCAC	1520
1521	AAATTTACAATTCCATACATCAATTGACGGAAGACCTATT CC T G C	1560
1561	AATCAGGGGAATTTTTTCAGCAACTATGAGTAGTGGGAGTA	1600
1601	ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTAC	1640
1641	TCCGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTA	1680
1681	AGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG	1720
1721	ATCGAATTGAATTTGTTCCGGCA	1743

FIGURE 2C

1	ATGGATAACAATCCGAACATCAATGAATGCATTTCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTTCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGGTGGGGATTTGATGCCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATCATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAGCGTGTATGGGGACCGGATTCTAGAGATTGGAT	680
	C G C T T	

FIGURE 3A

681	AAGATATAATCAATTTAGAAAGAGAATTAACACTAACTGTA	720
	T C C G C G G C C A T	
721	TTAGATATCGTTTCTCTATTTCCGAACATGATAGTAGAA	760
	G C T G C C CTCC	
761	CGTATCCAATTCGAACAGTTTCCCAATTAACAAGAGAAAT	800
	C C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTC	880
	T T T C A T C CTCC C C	
881	CACATTTGATGGATATACTTAATAGTATAACCATCTATAC	920
	C C CT G C C T C	
921	GGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA	960
	C C G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACATGCGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACCTTATATAGAAGACCTTTTAATATAG	1120
	C G T G C C C C	
1121	GGATAATAATCAACAACCTATCTGTTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGCTGAATTTAATAATATAATTCC	1400
	A T G C C C	

FIGURE 3B

1401	TTCATCACAAATTACACAAATACCTTTAACAAATCTACT	1440
	C T C C C A G C G	
1441	AATCTTGGCTCTGGAACCTTCTGTCGTTAAAGGACCAGGAT	1480
	C A G C	
1481	TTACAGGAGGAGATATTCTTCGAAGAACTTCACCTGGCCA	1520
	C T A T	
1521	GATTTCAACCTTAAGAGTAAATATTACTGCACCATTATCA	1560
	AGC C C T C C C T T	
1561	CAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCACAA	1600
	T C G T A A	
1601	ATTTACAATTCCATACATCAATTGACGGAAGACCTATTAA	1640
	C G C C C C G C	
1641	TCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTAAT	1680
	T C C C C TCA C C C C	
1681	TTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTC	1720
	G A C C A C C C	
1721	CGTTTAACTTTTCAAATGGATCAAGTGTATTTACGTTAAG	1760
	T C C T C C T C C C T	
1761	TGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAGAT	1800
	C G T T G C T C	
1801	CGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAG	1840
	T G G T C T C T	
1841	AATAT	1845
	G C	

FIGURE 3C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAGTGAATTGTTCCTGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCTTACAACCGCTATTCCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGGTGGGGATTGTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 4A

681	AAGGTATAATCAATTTAGAAAGAGAATTAACACTAACTGTA	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTGCTCTGTTCCTCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTTCGAACAGTTTCCCAATTAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACATATGGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACATCTGTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	

FIGURE 4B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAAATTCAGGACCAGGATTTA	1480
	C C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAATAA	1520
	A C C C C C C	
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACCTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTACCTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
	A TGCG	
1881	GCTGTTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	CTGT ACGTCTACA C AGCT G ACTC G CA TG	
1921	G 1921	

FIGURE 4C

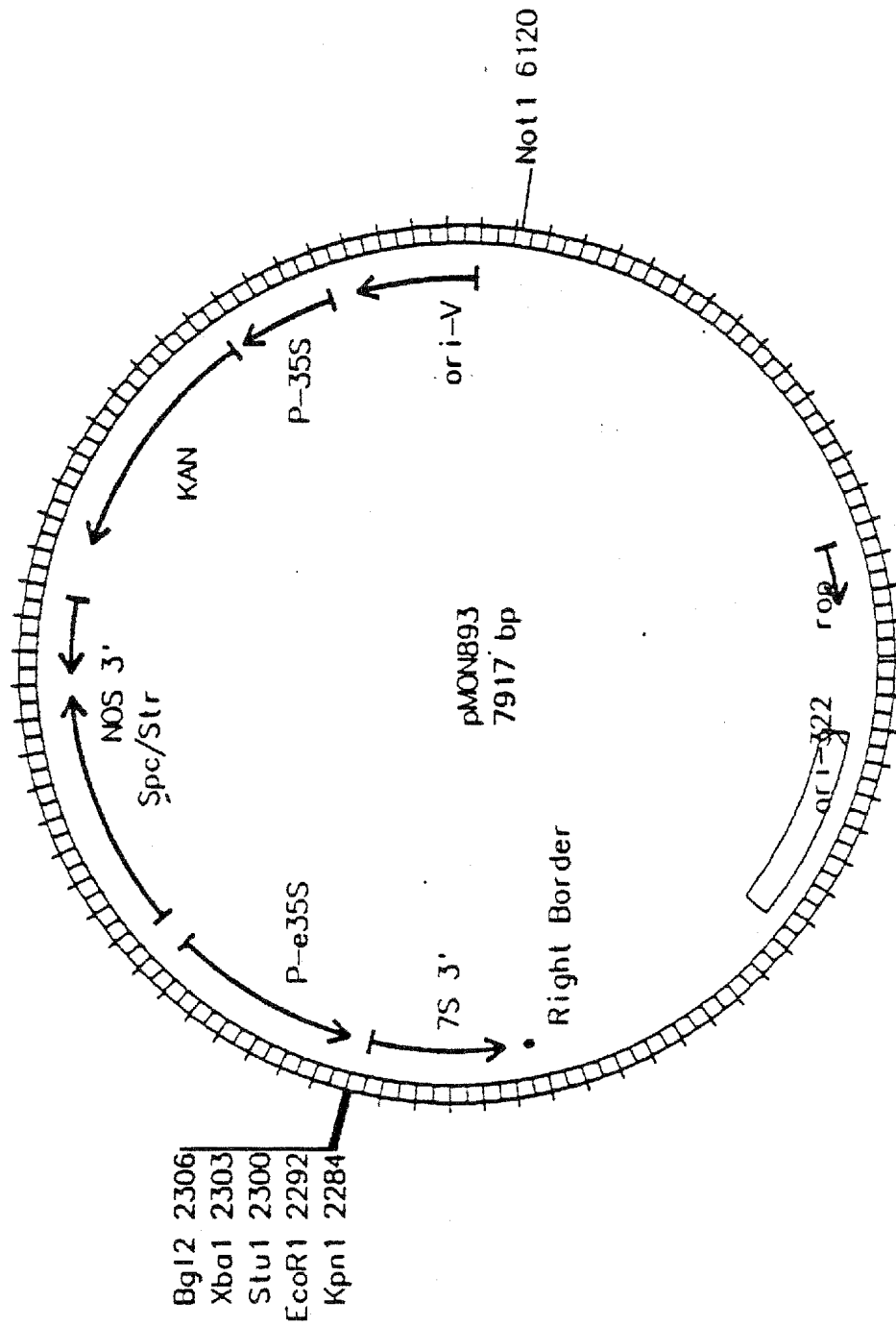


FIGURE 5

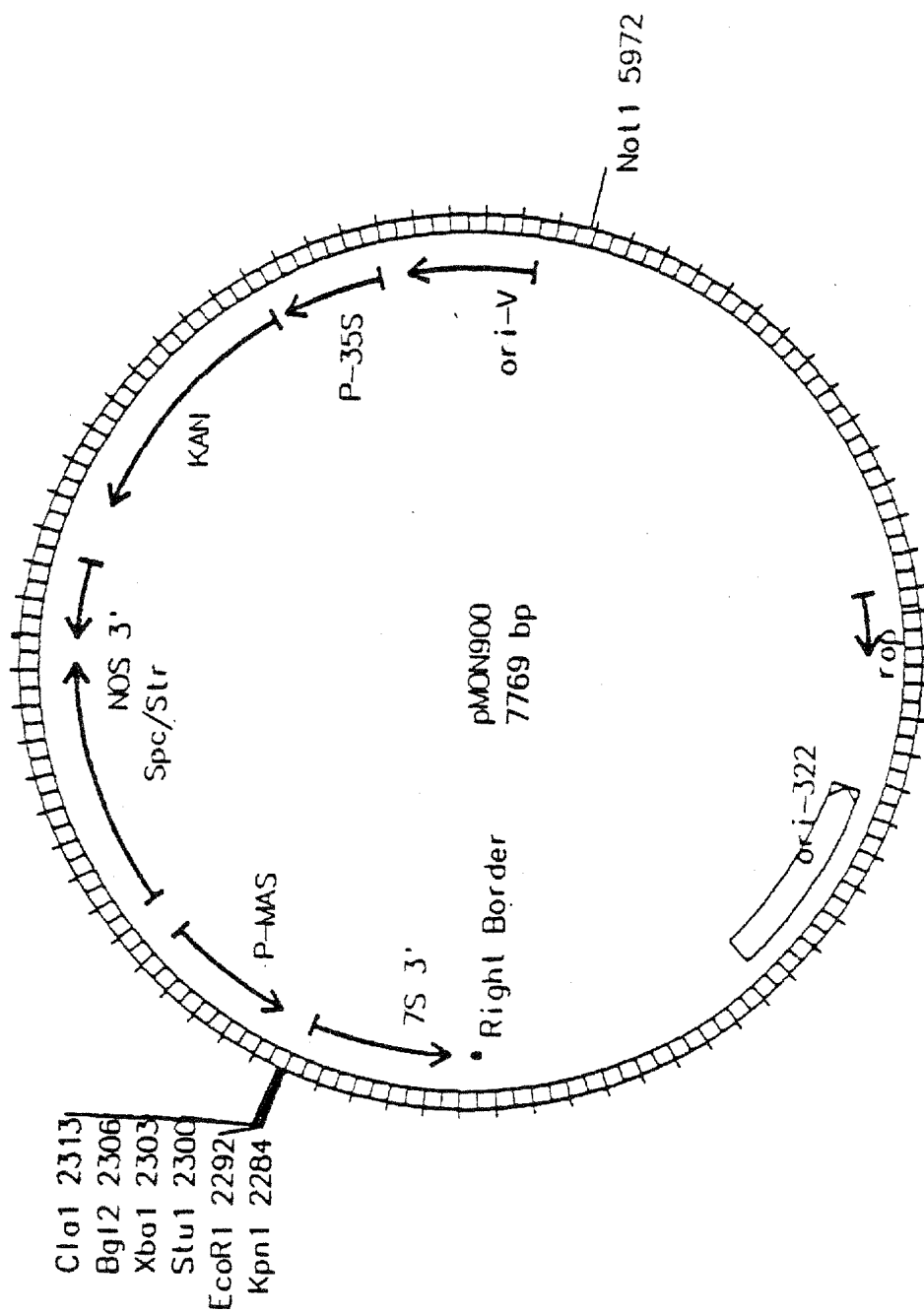


FIGURE 6

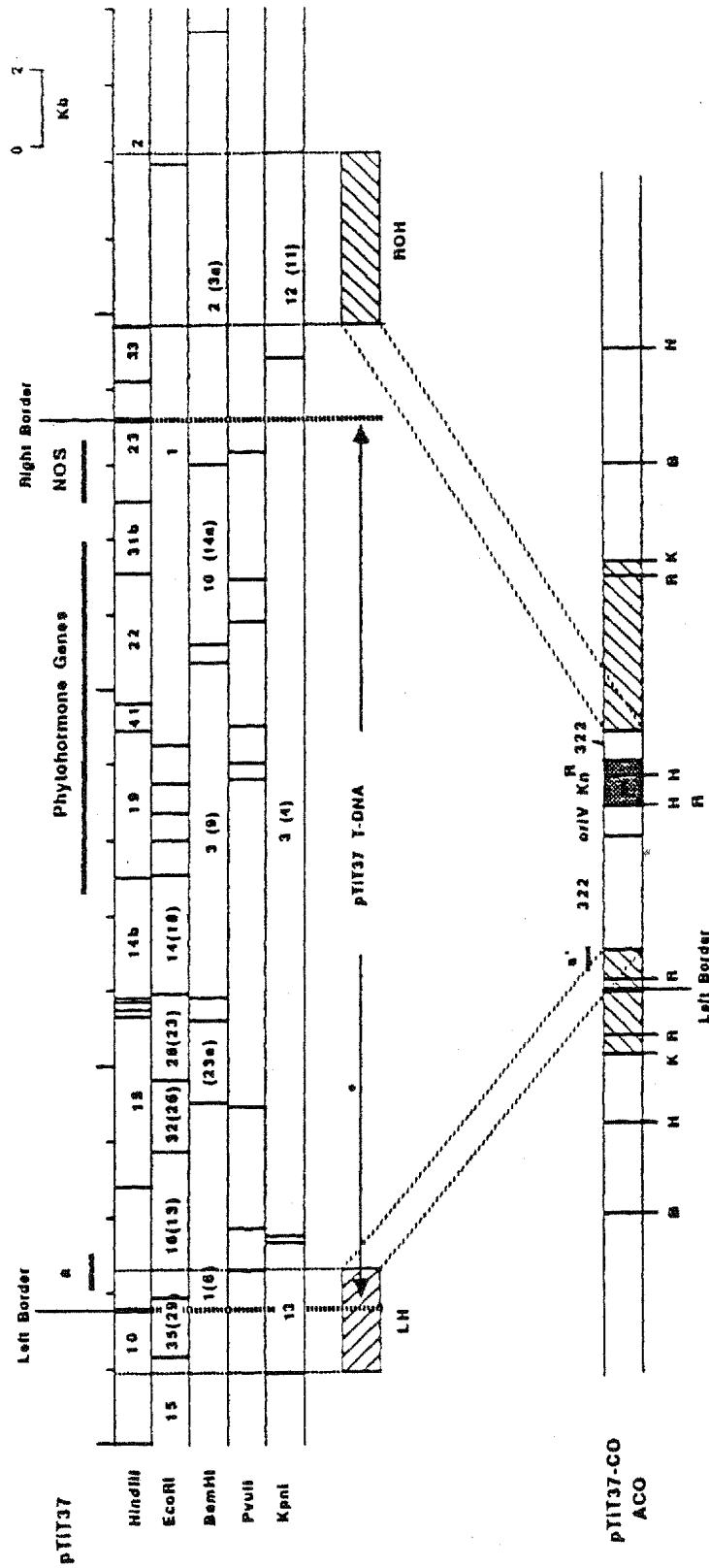


FIGURE 7

1	GAAAGAATAGAAACTGGTTACACCCCAATCGATATTTTCCT	40
	ATGGCC T C T C C C	
41	TGTCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGG	80
	CT G A G GC C C G C G A	
81	TGCTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGA	120
	G C TC C C C C T	
121	ATTTTGGTCCCTCTCAATGGGACGCATTCTTGTACAAA	160
	C A T C G G	
161	TTGAACAGTTAATTAACCAAGAATAGAAGAATTCGCTAG	200
	G G C G G C G C	
201	GAACCAAGCCATTCTAGATTAGAAGGACTAAGCAATCTT	240
	G C G G T G C	
241	TATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAG	280
	C C T GAGC C C	
281	ATCCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCA	320
	C TC CC C G A	
321	ATTCAATGACATGAACAGTGCCCTTACAACCGCTATTCCT	360
	C C T G C A C A	
361	CTTTTGCAGTTCAAATTATCAAGTTCCTCTTTTATCAG	400
	T G C C G C C C G C	
401	TATATGTTCAAGCTGCAAATTTACATTTATCAGTTTGTAG	440
	G C A T C T CC CAGC GC TC	
441	AGATGTTTCAGTGTTTGGACAAAGGTGGGGATTGTGATGCC	480
	C AGC G C T	
481	GCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTA	520
	A C C C C CC T G	
521	TTGGCAACTATACAGATTATGCTGTACGCTGGTACAATAC	560
	A C C CC C T T C	
561	GGGATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGG	600
	T C G G C T T	
601	GTAAGGTATAATCAATTTAGAAGAGAATTAACACTAACTG	640
	A T A C C G C G G C C A	
641	TATTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAG	680
	T G C T GT C C CTCC	

FIGURE 8A

681	AAGATATCCAATTTCGAACAGTTTCCCAATTAACAAGAGAA CC C T C T G C T C	720
721	ATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTT C T TC T G C C C C	760
761	TTCGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAG C T T T C A T C G CTCC C	800
801	TCCACATTTGATGGATATACTTAACAGTATAACCATCTAT C C C CT G C T C	840
841	ACGGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATC C C A AG G C T A C	880
881	AAATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATT G C C A T A CAGC C G	920
921	CACTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCA T C T C C C	960
961	CAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATA C T C C	1000
1001	GAACATTATCGTCCACTTTATATAGAAGACCTTTTAATAT C G T C G C C C	1040
1041	AGGGATAAATAATCAACAACCTATCTGTTCTTGACGGGACA C T C C C G T C A	1080
1081	GAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTG G C C T T C	1120
1121	TATACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAAT T G C T CT C	1160
1161	ACCGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTT C A C T C C	1200
1201	AGTCATCGATTAAGCCATGTTTCAATGTTTCGTTTCAGGCT TCC CA G G C G C C C A	1240
1241	TTAGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTT C C C TCC G C C C	1280
1281	CTCTTGGATACATCGTAGTGCTGAATTTAATAATATAATT C G C C C C C	1320
1321	GCATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAA C	1360
1361	ACTTTCCTTTTAATGGTTCTGTAATTTTCAGGACCAGGATT C C C C	1400

FIGURE 8B

1401	TACTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAAT	1440
	C A C C C C C C	
1441	AACATTCAGAAATAGAGGGTATATTGAAGTTCCAATTCCT	1480
1481	TCCCATCGACATCTACCAGATATCGAGTTCTGTGTACGGTA	1520
	C A GA	
1521	TGCTTCTGTAAACCCCGATTACCTCAACGTTAATTGGGGT	1560
	G T	
1561	AATTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTA	1600
	C C T	
1601	CGTCATTAGATAATCTACAATCAAGTGATTTTGGTTATTT	1640
	C C G C C C C C	
1641	TGAAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATA	1680
	C C C C	
1681	GTAGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAA	1720
	G C T	
1721	TAGACAGATTTGAATTTATTCCAGTTACTGCAACACTCGA	1760
	C C G C	
1761	GGCTGAA	1767
	G	

FIGURE 8C

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGGTGGGGATTTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 9A

681	AAGGTATAATCAATTTAGAAGAGAATTAACACTAACTGTÀ	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTTGCTCTGTCCCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTTCGAACAGTTTCCCAATTAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCACA	1040
	C T C C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACTATCTGTTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTÀ	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGGATACATCGTAGTGCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	

FIGURE 9B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAATTTAGGACCAGGATTTA	1480
	C C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAATAA	1520
	A C C C C C	
1521	CATTGAGAAATAGAGGGTATATTGAAGTTCCAATTCATTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTACCTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTCCTCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTTACGTCTACAAACCACTAGGGCTAAAAACAAAT	1920
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA	1960
1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTAA	2080
2081	ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT	2120

FIGURE 9C

2121	TACCATCCAAGGAGGGGATGACGTATTTAAAGAAAATTAC	2160
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
2281	GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT	2360
2361	TTCAGCCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCGTGTAGGGATGGAGAAAAGTGTGCCCATCATTGCGA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA	2560
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT	2680
2681	TGGAATGGGAAACAAATATCGTTTTATAAGAGGCAAAAGA	2720
2721	ATCTGTAGATGCTTTATTTGTAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800
2801	ATAAACGTGTTTCATAGCATTGAGAGCTTATCTGCCTGA	2840

FIGURE 9D

2841	GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGGAA	2880
2881	GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG	2920
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG	2960
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
3001	GAACAAAACAACCAACGTTCCGGTCCTTGTTGTTCCGGAAT	3040
3041	GGGAAGCAGAAAGTGTCAACAAGAAGTTCGTGTCTGTCCGGG	3080
3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGA	3120
3121	TATGGAGAAGGTTGCGTAACCATTTCATGAGATCGAGAACA	3160
3161	ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA	3200
3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
3241	GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA	3280
3281	ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATT	3320
3321	TGCGTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA	3360
3361	AGAGAGAATCCTTGTGAATTTAACAGAGGGTATAGGGATT	3400
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGAATTAC	3520
3521	TCCTTATGGAGGAA	3534

FIGURE 9E

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTGTTGGACAAAGGTGGGGATTTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	

FIGURE 10A

641	GATTAGAACGTGTAATGGGGACCGGATTCTAGAGATTGGGT C G G C T T A	680
681	AAGGTATAATCAATTTAGAAAGAGAATTAACACTAACTGTA T A C C G C G G C C A T	720
721	TTAGATATCGTTGCTCTGTTCCTCCGAATTATGATAGTAGAA G C T GT C C CTCC	760
761	GATATCCAATTTCGAACAGTTTCCCAATTAACAAGAGAAAT CC C T C T G C T C	800
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT C T TC T G C C C C C	840
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC T T T C A T C G CTCC C C	880
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC C C CT G C T C	920
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA C C A AG G C T A C G	960
961	ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCA C C A T A CAGC C G T	1000
1001	CTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCACA C T C C C C	1040
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA C T C C	1080
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAG C G T C G C C C C	1120
1121	GGATAAATAATCAACAACCTATCTGTTCTTGACGGGACAGA T C C C G T C A	1160
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGT G C C T T C T	1200
1201	TACAGAAAAAGCGGAACGGTAGATTGCTGGATGAAATAC G C T CT C C	1240
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG A C T C CTC	1280
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT C CA G G C G C C C A C	1320
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT C C TCC G C C C	1360

FIGURE 10B

1361	CTTGGATACATCGTAGTGCCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	
1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTCTCTGTAATTTTCAGGACCAGGATTTA	1480
	C C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAATAA	1520
	A C C C C C C	
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCCTTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTCACTCAACGTTAATTGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTCCTCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
1881	GCTGTTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	G C C C G C	
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA	1960
	G C G G	
1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
	C CC CAGC G C	
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080

FIGURE 10C

2081	ATAGGCAACCAGAACGTGGGTGGGGCGGAAGTACAGGGAT	2120
2121	TACCATCCAAGGAGGGGATGACGTATTTAAAGAAAATTAC G T C G C G G C	2160
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAAGCCTT CC C C G G C G C G G	2240
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
2281	GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG C C G CC C C	2320
2321	AAACAGTAAATGTGCCAGGTACGGGTTCCCTTATGGCCGCT	2360
2361	TTCAGCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
2441	GTTCTGTAGGGATGGAGAAAAGTGTGCCCATCATTGCGA	2480
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA	2560
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT G G	2680
2681	TGGAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA G C C C C	2720
2721	ATCTGTAGATGCTTTATTTGTAACTCTCAATATGATCAA	2760
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800

FIGURE 10D

2801	ATAAACGTGTTTCATAGCATTTCGAGAAGCTTATCTGCCTGA	2840
2841	GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGTAA	2880
2881	GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG	2920
	C C	
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG	2960
	C C C G C C C	
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
3001	GAACAAAACAACCAACGTTCCGGTCCTTGTTGTTCCGGAAT	3040
3041	GGGAAGCAGAAGTGTACACAAGAAGTTCGTGTCTGTCCGGG	3080
3081	TCGTGGCTATATCCTTCGTGTACAGCGTACAAGGAGGGG	3120
3121	TATGGAGAAGGTTGCGTAACCATTCATGAGATCGAGAACA	3160
3161	ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA	3200
3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
3241	GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA	3280
3281	ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATTA	3320
3321	TGCGTCAGTCTATGAAGAAAATCGTATACAGATGGACGA	3360
3361	AGAGAGAATCCTTGTGAATTTAACAGAGGGTATAGGGATT	3400
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
3481	GAAACGGAAGGAACATTTATCGTGACAGCGTGGAATTAC	3520
3521	TCCTTATGGAGGAA	3534

FIGURE 10E

1	ATGGATAACAATCCGAACATCAATGAATGCATTCCCTTATA	40
	C C A C A C	
41	ATTGTTTAAGTAACCCTGAAGTAGAAGTATTAGGTGGAGA	80
	C C G A T C T	
81	AAGAATAGAACTGGTTACACCCCAATCGATATTTCCCTTG	120
	C C T C T C C C	
121	TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTG	160
	CT G A G GC C C G C G A	
161	CTGGATTTGTGTTAGGACTAGTTGATATAATATGGGGAAT	200
	G C TC C C C C T	
201	TTTTGGTCCCTCTCAATGGGACGCATTTCTTGTAACAATT	240
	C A T C G G	
241	GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGA	280
	G G C G G C G C	
281	ACCAAGCCATTTCTAGATTAGAAGGACTAAGCAATCTTTA	320
	G C G G T G C	
321	TCAAATTTACGCAGAATCTTTTAGAGAGTGGAAGCAGAT	360
	C C T GAGC C C	
361	CCTACTAATCCAGCATTAAAGAGAAGAGATGCGTATTCAAT	400
	C TC CC C G A	
401	TCAATGACATGAACAGTGCCCTTACAACCGCTATTCCCTCT	440
	C C T G C A C AT	
441	TTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA	480
	G C C G C C C G C G	
481	TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAG	520
	C A T C T CC CAGC GC TC	
521	ATGTTTCAGTGTTTGGACAAAGGTGGGGATTGATGCCGC	560
	C AGC G C T	
561	GACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT	600
	A C C C C CC T G	
601	GGCAACTATACAGATTATGCTGTACGCTGGTACAATACGG	640
	A C C CC C T T C T	
641	GATTAGAACGTGTATGGGGACCGGATTCTAGAGATTGGGT	680
	C G G C T T A	

FIGURE 11A

681	AAGGTATAATCAATTTAGAAAGAGAATTAACACTAACTGTA	720
	T A C C G C G G C C A T	
721	TTAGATATCGTTGCTCTGTTCCCGAATTATGATAGTAGAA	760
	G C T GT C C CTCC	
761	GATATCCAATTCTGAACAGTTTCCCAATTAACAAGAGAAAT	800
	CC C T C T G C T C	
801	TTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT	840
	C T TC T G C C C C C	
841	CGAGGCTCGGCTCAGGGCATAGAAAGAAGTATTAGGAGTC	880
	T T T C A T C G CTCC C C	
881	CACATTTGATGGATATACTTAACAGTATAACCATCTATAC	920
	C C CT G C T C	
921	GGATGCTCATAGGGGTTATTATTATTGGTCAGGGCATCAA	960
	C C A AG G C T A C G	
961	ATAATGGCTTCTCCTGTAGGGTTTTTCGGGGCCAGAATTCA	1000
	C C A T A CAGC C G T	
1001	CTTTTCCGCTATATGGAACCTATGGGAAATGCAGCTCCACA	1040
	C T C C C	
1041	ACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA	1080
	C T C C	
1081	ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAG	1120
	C G T C G C C C C	
1121	GGATAAATAATCAACAACCTATCTGTTCTTGACGGGACAGA	1160
	T C C C G T C A	
1161	ATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA	1200
	G C C T T C T	
1201	TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATAC	1240
	G C T CT C C	
1241	CGCCACAGAATAACAACGTGCCACCTAGGCAAGGATTTAG	1280
	A C T C CTC	
1281	TCATCGATTAAAGCCATGTTTCAATGTTTCGTTTCAGGCTTT	1320
	C CA G G C G C C C A C	
1321	AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCT	1360
	C C TCC G C C C	
1361	CTTGATACATCGTAGTGCTGAATTTAATAATATAATTGC	1400
	C G C C C C C	

FIGURE 11B

1401	ATCGGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC	1440
	C	
1441	TTTCTTTTAAATGGTTCTGTAAATTCAGGACCAGGATTTA	1480
	C C C C	
1481	CTGGTGGGGACTTAGTTAGATTAAATAGTAGTGGAATAA	1520
	A C C C C C	
1521	CATTCAGAATAGAGGGTATATTGAAGTTCCAATTCACTTC	1560
1561	CCATCGACATCTACCAGATATCGAGTTCGTGTACGGTATG	1600
	C A GA	
1601	CTTCTGTAACCCCGATTACCTCAACGTTAATTGGGGGTAA	1640
	G T	
1641	TTCATCCATTTTTTCCAATACAGTACCAGCTACAGCTACG	1680
	C C T C	
1681	TCATTAGATAATCTACAATCAAGTGATTTTGGTTATTTTG	1720
	C G C C C C	
1721	AAAGTGCCAATGCTTTTACATCTTCATTAGGTAATATAGT	1760
	C C C C	
1761	AGGTGTTAGAAATTTTAGTGGGACTGCAGGAGTGATAATA	1800
	G C T C	
1801	GACAGATTTGAATTTATTCCAGTTACTGCAACACTCGAGG	1840
	C G C	
1841	CTGAATATAATCTGGAAAGAGCGCAGAAGGCGGTGAATGC	1880
	G C C T G C T C	
1881	GCTGTTTACGTCTACAAACCAACTAGGGCTAAAAACAAAT	1920
	C C C C C T G T CT G T C	
1921	GTAACGGATTATCATATTGATCAAGTGTCCAATTTAGTTA	1960
	T T C C C C G C	
1961	CGTATTTATCGGATGAATTTTGTCTGGATGAAAAGCGAGA	2000
	C CC TAGC G C C C C G T	
2001	ATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGAT	2040
	C C T C C T C C	
2041	GAACGCAATTTACTCCAAGATTCAAATTTCAAAGACATTA	2080
	GA G C CT G C C C C	
2081	ATAGGCAACCAGAACSTGGGTGGGGCGGAAGTACAGGGAT	2120
	C G T T C C	

FIGURE 11C

2121	TACCATCCAAGGAGGGGATGACCTATTTAAAGAAAATTAC	2160
	C C C T G C G G C	
2161	GTCACACTATCAGGTACCTTTGATGAGTGCTATCCAACAT	2200
	C C C A T C C C T C	
2201	ATTTGTATCAAAAAATCGATGAATCAAAATTAAAGCCTT	2240
	C C G G G C C C	
2241	TACCCGTTATCAATTAAGAGGGTATATCGAAGATAGTCAA	2280
	C A G C T C C C C	
2281	GACTTAGAAATCTATTTAATTCGCTACAATGCAAAACATG	2320
	C T C CG CA G C G C	
2321	AAACAGTAAATGTGCCAGGTACGGGTCCTTATGGCCGCT	2360
	G C G C T C C A	
2361	TTCAGCCCAAGTCCAATCGGAAAGTGTGGAGAGCCGAAT	2400
	T TC C T G T C	
2401	CGATGCGCGCCACACCTTGAATGGAATCCTGACTTAGATT	2440
	A T G G C	
2441	GTTCTGTAGGGATGGAGAAAAGTGTGCCCATCATTGCA	2480
	C C C C G C T	
2481	TCATTTCTCCTTAGACATTGATGTAGGATGTACAGACTTA	2520
	C G C G T C G	
2521	AATGAGGACCTAGGTGTATGGGTGATCTTTAAGATTAAGA	2560
	C A C C C C	
2561	CGCAAGATGGGCACGCAAGACTAGGGAATCTAGAGTTTCT	2600
	C C A T C C T	
2601	CGAAGAGAAACCATTAGTAGGAGAAGCGCTAGCTCGTGTG	2640
	G C T T C	
2641	AAAAGAGCGGAGAAAAAATGGAGAGACAAACGTGAAAAAT	2680
	G A G G G G C	
2681	TGGAATGGGAAACAAATATCGTTTATAAAGAGGCAAAAGA	2720
	C T C C G C	
2721	ATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAA	2760
	G C G G C G C G	
2761	TTACAAGCGGATACGAATATTGCCATGATTCATGCGGCAG	2800
	G C C C C C C C C	
2801	ATAAACGTGTTTCATAGCATTGCGAGAAGCTTATCTGCCTGA	2840
	C G C T G CT	

FIGURE 11D

2841	GCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTGA	2880
	T C C T G C T C C C G	
2881	GAATTAGAAGGGCGTATTTTCACTGCATTCTCCCTATATG	2920
	C T G A C T C T G C	
2921	ATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATGG	2960
	C C C G C C C C	
2961	CTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAA	3000
	C CAG T T G C G G	
3001	GAACAAAACAACCAACGTTCCGGTCCTTGTTGTTCCGGAAT	3040
	G T G C G G T G	
3041	GGGAAGCAGAAGTGTACACAAGAAGTTCGTGTCTGTCCGGG	3080
	T C G A A A	
3081	TCGTGGCTATATCCTTCGTGTACACAGCGTACAAGGAGGGA	3120
	A A C T C G C T	
3121	TATGGAGAAGGTTGCGTAACCATTTCATGAGATCGAGAACA	3160
	C T G G C C	
3161	ATACAGACGAACTGAAGTTTAGCAACTGCGTAGAAGAGGA	3200
	C C G T CTC C G A	
3201	AATCTATCCAAATAACACGGTAACGTGTAATGATTATACT	3240
	C C C T T C C C C	
3241	GTAAATCAAGAAGAATACGGAGGTGCGTACACTTCTCGTA	3280
	G G G C AGC	
3281	ATCGAGGATATAACGAAGCTCCTTCCGTACCAGCTGATTA	3320
	CA T C T T C	
3321	TGCGTCAGTCTATGAAGAAAAATCGTATACAGATGGACGA	3360
	C C G C G G C C CA	
3361	AGAGAGAATCCTTGTGAATTTAACAGAGGGTATAGGGATT	3400
	C T C C G C T C C	
3401	ACACGCCACTACCAGTTGGTTATGTGACAAAAGAATTAGA	3440
	A T C T C G GC T	
3441	ATACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGA	3480
	G T T G C A G C C T	
3481	GAAACGGAAGGAACATTTATCGTGGACAGCGTGGAATTAC	3520
	C G C C GC T	
3521	TCCTTATGGAGGAA	3534
	T G	

FIGURE 11E

1	ATGACTGCAGATAATAATACGGAAGCACTAGATAGCTCTA	40
	C C C C C C C T	
41	CAACAAAAGATGTCATTCAAAAAGGCATTTCGGTAGTAGG	80
	C T G T C G G T C T G	
81	TGATCTCCTAGGCGTAGTAGGTTTCCCGTTTGGTGGAGCG	120
	A C T G G T A T C C C	
121	CTTGTTTCGTTTTATACAAACTTTTTAAATACTATTTGGC	160
	C GAGC C C C C C	
161	CAAGTGAAGACCCGTGGAAGGCTTTTATGGAACAAGTAGA	200
	C G T A A C G T	
201	AGCATTGATGGATCAGAAAATAGCTGATTATGCAAAAAAT	240
	TC T G T A C G C	
241	AAAGCTCTTGACAGAGTTACAGGGCCTTCAAAATAATGTGG	280
	G T G A C C G C G	
281	AAGATTATGTGAGTGCATTGAGTTCATGGCAAAAAAATCC	320
	G C C TCCAGC G G C	
321	TGTGAGTTCACGAAATCCACATAGCCAGGGGCGGATAAGA	360
	T C CA T C A TA C	
361	GAGCTGTTTTCTCAAGCAGAAAGTCATTTTCGTAATTCAA	400
	T C C TCC C CA A C	
401	TGCCTTCGTTTGCAATTTCTGGATACGAGGTTCTATTTCT	440
	AGC T C C T T C	
441	AACAACATATGCACAAGCTGCCAACACACATTTATTTTAA	480
	C T C T C C G C C	
481	CTAAAAGACGCTCAAATTTATGGAGAAGAATGGGGATACG	520
	T G C G	
521	AAAAAGAAGATATTGCTGAATTTTATAAAAGACAACATAAA	560
	G G C G C C GC T T	
561	ACTTACGCAAGAATATACTGACCATTGTGTCAAATGGTAT	600
	G C C G C C G	
601	AATGTTGGATTAGATAAATTAAGAGGTTTATCTTATGAAT	640
	C TC C GC C C T C C G	
641	CTTGGGTAAACTTTAACCGTTATCGCAGAGAGATGACATT	680
	G C A A CA G C	

FIGURE 12A

681	AACAGTATTAGATTTAATTGCACTATTTCCATTGTATGAT	720
	G T GC C C T C C C C	
721	GTTCCGGCTATACCCAAAAGAGTTAAAACCGAATTAACAA	760
	GA A C G G T GC T C	
761	GAGACGTTTTTAACAGATCCAATTGTCGGAGTCAACAACCT	800
	GC C T C T	
801	TAGGGGCTATGGAACAACCTTCTCTAATATAGAAAATTAT	840
	T T AGC C C C	
841	ATTCGAAAACACATCTATTTGACTATCTGCATAGAATTC	880
	A G C C T C	
881	AATTTACACGCGGTTCCAACCAGGATATTATGGAAATGA	920
	C AA T C T C	
921	CTCTTTCAATTATTGGTCCGGTAATTATGTTTCAACTAGA	960
	C C C C C	
961	CCAAGCATAGGATCAAATGATATAATCACATCTCCATTCT	1000
	T T C C C	
1001	ATGGAAATAAATCCAGTGAACCTGTACAAAATTTAGAATT	1040
	T C G G G CC T G	
1041	TAATGGAGAAAAAGTCTATAGAGCCGTAGCAAATACAAAT	1080
	C C C G C C C	
1081	CTTGCGGTCTGGCCGTCCGCTGTATATTCAGGTGTTACAA	1120
	C T G A A T C C C	
1121	AAGTGAATTTAGCCAATATAATGATCAAACAGATGAAGC	1160
	G G T G C G C G	
1161	AAGTACACAAACGTACGACTCAAAAAGAAATGTTGGCGCG	1200
	C C C G T C C T C A	
1201	GTCAGCTGGGATTCTATCGATCAATTGCCTCCAGAAACAA	1240
	TCT C C	
1241	CAGATGAACCTCTAGAAAAGGGATATAGCCATCAACTCAA	1280
	C AT G G C C C T	
1281	TTATGTAATGTGCTTTTTAATGCAGGGTAGTAGAGGAACA	1320
	C G C G A TCC G C	
1321	ATCCCAGTGTTAACTTGGACACATAAAAGTGTAGACTTTT	1360
	T G C C GTCC G C	
1361	TTAACATGATTGATTCGAAAAAATTACACAACCTCCGTT	1400
	C C AGC G G C T C	

FIGURE 12B

1401	AGTAAAGGCATATAAGTTACAATCTGGTGCTTCCGTTGTC	1440
	G G A C C C G	
1441	GCAGGTCCTAGGTTTACAGGAGGAGATATCATTCAATGCA	1480
	C A C T T C C G	
1481	CAGAAAATGGAAGTGCGGCAACTATTTACGTTACACCGGA	1520
	G C C C A T C G T	
1521	TGTGTCGTA CTCTCAAAAATATCGAGCTAGAATTCATTAT	1560
	T G G CA G AC T C	
1561	GCTTCTACATCTCAGATAACATTTACACTCAGTTTAGACG	1600
	A CAGC C C C C G T	
1601	GGGCACCATTTTAATCAATACTATTTTCGATAAAAACGATAAA	1640
	A C C C G T C T C G C C	
1641	TAAAGGAGACACATTAACGTATAATTCATTTAATTTAGCA	1680
	C T TC C A C AGC C C G	
1681	AGTTTCAGCACACCATTTCGAATTATCAGGGAATAACTTAC	1720
	T C C C C TC T	
1721	AAATAGGCGTCCACAGGATTAAGTGCTGGAGATAAAGTTTA	1760
	G C C TC C C C C C	
1761	TATAGACAAAATTGAATTTATTCCAGTGAAT	1791
	C C G G C C C	

FIGURE 12C

1	ATG AATAATGTATTGAATAGTGGGAAGAACAACTATTT	40
	GAC C C C CTC T C C	
41	GTGATGCGTATAATGTAGTAGCCCATGATCCATTTAGTTT	80
	C C A C C C G T C C C	
81	TGAACATAAATCATTAGATACCATCCAAAAAGAATGGATG	120
	C C GAGCC C C T T G G G	
121	GAGTGGAAAAGAACAGATCATAGTTTATATGTAGCTCCTG	160
	A C T T C CTC C C C C A	
161	TAGTCGGAACGTGTGTCTAGTTTTTTGCTAAAGAAAGTGGG	200
	G T A C C CC T C G C	
201	GAGTCTTATTGGAAAAAGGATATTGAGTGAATTATGGGGG	240
	CTC C C C T C TCC C C T	
241	ATAATATTTCTAGTGGTAGTACAAATCTAATGCAAGATA	280
	C C ATC GTCC T C C	
281	TTTTAAGGGAGACAGAACAATTCCTAAATCAAAGACTTAA	320
	C G C G T C C GC T C	
321	TACAGATACCCTTGCTCGTGTAATGCAGAATTGATAGGG	360
	C T T G A A C C T G C T	
361	CTCCAAGCGAATATAAGGGAGTTTAATCAACAAGTAGATA	400
	A C TC T C C G G C	
401	ATTTTTTAAACCCTACTCAAAACCCTGTTCTCTTTATCAAT	440
	C C G T A G T G C T C	
441	AACTTCTTCGGTTAATACAATGCAGCAATTATTTCTAAAT	480
	C C G C T C C C C C	
481	AGATTACCCCAGTTCCAGATACAAGGATACCAGTTGTTAT	520
	G T T T C C CC	
521	TATTACCTTTATTTGCACAGGCAGCCAATATGCATCTTTC	560
	TC T AC C T T C CT G	
561	TTTTATTAGAGATGTTATTCTTAATGCAGATGAATGGGGT	600
	C C AC T C G C C C T C A	
601	ATTCAGCAGCAACATTACGTACGTATCGAGATTACCTGA	640
	C T C TC TA G A CA C T	
641	GAAATTATACAAGAGATTATTCTAATTATTGTATAAATAC	680
	G C C TC T C C C C C C	

FIGURE 13A

681	GTATCAAAC [·] TGCGTTT [·] AGAGGGTTAAACACCCGTTTACAC [·]	720
	T G C C T AC C T TA GC T	
721	GATATGTTAGAATTTAGAACATATATGTTTTTAAATGTAT [·]	760
	C C T G C G C C CC T C G	
761	TTGAATATGTATCCATTTGGTCAATTGTTTAAATATCAGAG [·]	800
	G C CAG AGTC C C G C	
801	TCTTATGGTATCTTCTGGCGCTAATTTATATGCTAGCGGT [·]	840
	CT G G C A C C C C CTCT C	
841	AGTGGACCACAGCAGACACAATCATTTCAGCACAAAAC [·]	880
	A T GAGC C T G	
881	GGCCATTTTTATATTCTCTTTTCCAAGTTAATTCGAATTA [·]	920
	C G AGCT G C C C C	
921	TATATTATCTGGTATTAGTGGTACTAGGCTTCTATTACC [·]	960
	C TC CAG CTC G C A C C A	
961	TTCCCTAATATTGGTGGTTTACCGGGTAGTACTACAAC [·] TC	1000
	T C C AC T A CTCC C	
1001	ATTCATTGAATAGTGCCAGGGTTAATTATAGCGGAGGAGT [·]	1040
	AGCC T CTC A G C C T T	
1041	TTCATCTGGTCTCATAGGGGCGACTAATCTCAATCACAAC [·]	1080
	CAGC AT G T T A CT G C	
1081	TTTAATTGCAGCACGGTCCTCCCTCCTTTATCAACACCAT [·]	1120
	C TC C T G A C GAGC G	
1121	TTGTTAGAAGTTGGCTGGATTCAAGTACAGATCGAGAGGG [·]	1160
	G GTCC T CAGC T C A	
1161	CGTTGCTACCTCTACGAATTGGCAGACAGAATCCTTTCAA [·]	1200
	A A C A C G C	
1201	ACAACTTTAAGTTTAAGGTGTGGTGCTTTTTCAGCCCGTG [·]	1240
	C C T CC TC A C T A	
1241	GAAATTCAAAC [·] TATTTCCAGATTATTTTATCCGTAATAT [·]	1280
	G C T C C C TA G C	
1281	TTCTGGGGTTCCTTTAGTTATTAGAAACGAAGATCTAACA [·]	1320
	C T C C C C G T C C C	
1321	AGACCGTTACACTATAACCAAATAAGAAATATAGAAAGTC [·]	1360
	C T AC T T C G T G C C GTC	
1361	CTTCGGGAACACCTGGTGGAGCACGGGCCTATTTGGTATC [·]	1400
	A C T T A A T A A T CC C G	

FIGURE 13B

1401	TGTGCATAACAGAAAAAATAATATCTATGCCGCTAATGAA	1440
	C G G C C C T C C G	
1441	AATGGTACTATGATCCATTGCGGCCAGAGATTATACAG	1480
	C C T CC T A C T	
1481	GATTTACTATATCGCCAATACATGCCACTCAAGTGAATAA	1520
	C C C T C T C C	
1521	TCAAACCTCGAACATTTATTTCTGAAAAATTTGGAAATCAA	1560
	G A C C C C C G C	
1561	GGTGATTCCTTAAGATTTGAACAAAGCAACACGACAGCTC	1600
	C G G C G TC T C A	
1601	GTTATACGCTTAGAGGGAATGGAAATAGTTACAATCTTTA	1640
	G C TT G C C C C	
1641	TTTAAGAGTATCTTCAATAGGAAATTCAACTATTTCGAGTT	1680
	C G TAGC C T T C C C C T	
1681	ACTATAAACGGTAGAGTTTATACTGTTTCAAATGTTAATA	1720
	C C AC T C A C T G C	
1721	CCACTACAAATAACGATGGAGTTAATGATAATGGAGCTCG	1760
	T A G C T C C C C CA	
1761	TTTTTCAGATATTAATATCGGTAATATAGTAGCAAGTGAT	1800
	A CAGC C C C T C C C G CTC C	
1801	AATACTAATGTAACGCTAGATATAAATGTGACATTAAACT	1840
	C C T TT G C C CC C T	
1841	CCGGTACTCCATTTGATCTCATGAATATTATGTTTGTGCC	1880
	T A C C	
1881	AACTAATCTTCCACCACTTTAT	1902
	C C T T G C	

FIGURE 13C

1	ATGGAGGAAAATAATCAAAATCAATGCATACCTTACAATT	40
	G C C C T A C	
41	GTTTAAGTAATCCTGAAGAAGTACTTTTGGATGGAGAACG	80
	C G C A G T G C T	
81	GATATCAACTGGTAATTCATCAATTGATATTTCTCTGTCA	120
	C T C C T C C C C C T C	
121	CTTGTTTCAGTTTCTGGTATCTAACTTTGTACCAGGGGGAG	160
	T G C C A G C C G T T	
161	GATTTTTAGTTGGATTAATAGATTTTGTATGGGGAATAGT	200
	G C C T C C T C C C T C	
201	TGGCCCTTCTCAATGGGATGCATTTCTAGTACAAATTGAA	240
	T A C G G G	
241	CAATTAATTAATGAAAGAATAGCTGAATTTGCTAGGAATG	280
	G G C C G G C G C C C	
281	CTGCTATTGCTAATTTAGAAGGATTAGGAAACAATTTCAA	320
	C C C G G C T C	
321	TATATATGTGGAAGCATTAAAGAATGGGAAGAAGATCCT	360
	C C G C C G G C	
361	AATAATCCAGAAACCAGGACCAGAGTAATTGATCGCTTTC	400
	C G C C T G G C C A A C A	
401	GTATACTTGATGGGCTACTTGAAAGGGACATTCCTTCGTT	440
	A C T G C C C T G G A T C A C	
441	TCGAATTTCTGGATTTGAAGTACCCCTTTTATCCGTTTAT	480
	C A C C C T T C G G C	
481	GCTCAAGCGGCCAATCTGCATCTAGCTATATTAAGAGATT	520
	A T T C C C C T C A	
521	CTGTAATTTTGGAGAAAGATGGGGATTGACAACGATAAA	560
	G C C G G C T C	
561	TGTCAATGAAAACCTATAATAGACTAATTAGGCATATTGAT	600
	C G T C C T C C C	
601	GAATATGCTGATCACTGTGCAAATACGTATAATCGGGGAT	640
	G C C C T C C C C T C	
641	TAAATAATTTACCGAAATCTACGTATCAAGATTGGATAAC	680
	G C C C T G T T	
681	ATATAATCGATTACGGAGAGACTTAACATTGACTGTATTA	720
	C C C A G G A G C C C A T G	

FIGURE 14A

721	GATATCGCCGCTTTCTTTCCAAACTATGACAAATAGGAGAT	760
	C T A C G C	
761	ATCCAATTCAGCCAGTTGGTCAACTAACAAGGGAAGTTTA	800
	C T C A G T C A C	
801	TACGGACCCATTAATTAATTTAATCCACAGTTACAGTCT	840
	T C T C C C T G AAG	
841	GTAGCTCAATTACCTACTTTTAACGTTATGGAGAGCAGCC	880
	C C C T C A C C TC	
881	GAATTAGAAATCCTCATTATTTGATATATTGAATAATCT	920
	T C G C A C G C C C C	
921	TACAATCTTTACGGATTGGTTTAGTGTTGGACGCAATTTT	960
	T C C C C G T C C	
961	TATTGGGGAGGACATCGAGTAATATCTAGCCTTATAGGAG	1000
	T C A G C C CTCT T	
1001	GTGGTAACATAACATCTCCTATATATGGAAGAGAGGCGAA	1040
	G T C C C T A	
1041	CCAGGAGCCTCCAAGATCCTTTACTTTTAATGGACCGGTÀ	1080
	A C TAGT G C C C T A C	
1081	TTTAGGACTTTATCAAATCCTACTTTACGATTATTACAGC	1120
	C A C G T C C GA GC C .	
1121	AACCTTGGCCAGCGCCACCATTTAATTTACGTGGTGTGA	1160
	T T C CC TA A	
1161	AGGAGTAGAATTTTCTACACCTACAAATAGCTTTACGTAT	1200
	G C T G C T C CTC C T C	
1201	CGAGGAAGAGGTACGGTTGATTCTTTAACTGAATTACCGC	1240
	A T A C C G C C C A	
1241	CTGAGGATAATAGTGTGCCACCTCGCGAAGGATATAGTCA	1280
	A C C CA G C CTCC	
1281	TCGTTTATGTCATGCAACTTTTGTTCAAAGATCTGGAACA	1320
	CA G G C C C C G GC T C T	
1321	CCTTTTTTAACAACCTGGTGTAGTATTTTCTTGGACCGATC	1360
	A CC C T A A T G C A T	
1361	GTAGTGCAACTCTTACAAATACAATTGATCCAGAGAGAAT	1400
	T C T C C G	

FIGURE 14B

1401	TAATCAAATACCTTTAGTGAAAGGATTTAGAGTTTGGGGG	1440
	C C A G C G T C C T G A	
1441	GGCACCTCTGTCATTACAGGACCAGGATTTACAGGAGGGG	1480
	A T C C C T	
1481	ATATCCTTCGAAGAAATACCTTTGGTGATTTTGTATCTCT	1520
	T A C T C C GAGC	
1521	ACAAGTCAATATTAATTCACCAATTACCCAAAGATACCGT	1560
	C T C C C C T T T	
1561	TTAAGATTTGTTACGCTTCCAGTAGGGATGCACGAGTTA	1600
	C C G A T T C C C T C T A C	
1601	TAGTATTAACAGGAGCGGCATCCACAGGAGTGGGAGGCCA	1640
	C G C C C C A T T C T C T A	
1641	AGTTAGTGTAATATGCCTCTTCAGAAAATATGGAAATA	1680
	CTCC G C A C G G C	
1681	GGGGAGAACTTAACATCTAGAACATTTAGATATACCGATT	1720
	C G C G C C C C	
1721	TTAGTAATCCTTTTTTCATTTAGAGCTAATCCAGATATAAT	1760
	CTC C CAGT C C T C C T C C	
1761	TGGGATAAGTGAACAACCTCTATTTGGTGCAGGTTCTATT	1800
	C T C C A T AGC C	
1801	AGTAGCGGTGAACCTTTATATAGATAAAATTGAAATTATTC	1840
	TCATCT C T G C T C G G C	
1841	TAGCAGATGCAACATTTGAAGCAGAATCTGATTTAGAAAG	1880
	T C C T C C C G T G A C A C C T G	
1881	AGCACAAAAGGCGGTGAATGCCCTGTTTACTTCTTCCAAT	1920
	C G T C C C C A	
1921	CAAATCGGGTTAAAAACCGATGTGACGGATTATCATATTG	1960
	GC T C G T A C T T C C	
1961	ATCAAGTATCCAATTTAGTGGATTGTTTATCAGATGAATT	2000
	C G C G CACC ACC TAGC G	
2001	TTGTCTGGATGAAAAGCGAGAATTGTCCGAGAAAGTCAAA	2040
	C C C C G T C C T	
2041	CATGCGAAGCGACTCAGTGATGAGCGGAATTTACTTCAAG	2080
	C C T C C A C C T G	
2081	ATCCAAACTTCAGAGGGATCAATAGACAACCAGACCGTGG	2120
	CT C A AC C G G A	

FIGURE 14C

2121	CTGGAGAGGAAGTACAGATATTACCATCCAAGGAGGAGAT	2160
	T G T C C GG C C C	
2161	GACGTATTCAAAGAGAATTACGTCACTACCGGGTACCG	2200
	T G G C C CT C A TT	
2201	TTGATGAGTGCTATCCAACGTATTTATATCAGAAAATAGA	2240
	C C C T C C G C G C	
2241	TGAGTCGAAATTAAAAGCTTATACCCGTTATGAATTAAGA	2280
	C C C C TC A G C C T	
2281	GGGTATATCGAAGATAGTCAAGACTTAGAAATCTATTTGA	2320
	C C C C C T C C	
2321	TCCGTTACAATGCAAAACACGAAATAGTAAATGTGCCAGG	2360
	A G C G G CC G C	
2361	CACGGGTTCCCTTATGGCCGCTTTCAGCCCAAATGCCAATC	2400
	T T C C A T TCT C T	
2401	GGAAAGTGTGGAGAACCGAATCGATGCGCGCCACACCTTG	2440
	G G T CA T	
2441	AATGGAATCCTGATCTAGATTGTTCCCTGCAGAGACGGGGA	2480
	G CT G C C G T C	
2481	AAAATGTGCACATCATTCCCATCATTTACCTTGGATATT	2520
	G G C C T C T C C	
2521	GATGTTGGATGTACAGACTTAAATGAGGACTTAGGTGTAT	2560
	G T C G C C A C	
2561	GGGTGATATTCAAGATTAAGACGCAAGATGGCCATGCAAG	2600
	C C C C C A C	
2601	ACTAGGGAATCTAGAGTTTCTCGAAGAGAAACCATTATTA	2640
	T C C T GG C	
2641	GGGGAAGCACTAGCTCGTGTGAAAAGAGCGGAGAAGAAGT	2680
	T T C G A	
2681	GGAGAGACAAACGAGAGAAACTGCAGTTGGAAACAAATAT	2720
	G T CG A G T C	
2721	TGTTTATAAAGAGGCCAAAAGAATCTGTAGATGCTTTATTT	2760
	C C G C G C G G C	
2761	GTAAACTCTCAATATGATAGATTACAAGTGGATACGAACA	2800
	G C CAG G CC C C	
2801	TCGCCATGATTCATGCGGCAGATAAACCGGTTTCATAGAAT	2840
	C C C C T G C C	

FIGURE 14D

2841	CCGGGAAGCGTATCTGCCAGAGTTGTCTGTGATTCCAGGT	2880
	T T G T CT T C C T	
2881	GTCAATGCGGCCATTTTCGAAGAATTAGAGGGACGTATTT	2920
	G C T C G C T C	
2921	TTACAGCGTATTCCTTATATGATGCGAGAAATGTCATTAA	2960
	C A TC G C C C C	
2961	AAATGGCGATTTCAATAATGGCTTATTATGCTGGAACGTG	3000
	G C T C C C CAGC T	
3001	AAAGGTCATGTAGATGTAGAAGAGCAAAACAACCACCGTT	3040
	G C G G A G T G	
3041	CGGTCCTTGTTATCCCAGAATGGGAGGCAGAAGTGTCACA	3080
	C G G G T G A T C	
3081	AGAGGTTCTGTCTGTCCAGGTCGTGGCTATATCCTTCGT	3120
	A A A A C T C	
3121	GTCACAGCATATAAAGAGGGATATGGAGAGGGCTGCGTAA	3160
	G C T C G C T T G	
3161	CGATCCATGAGATCGAAGACAATACAGACGAACTGAAATT	3200
	C C GA C C G T G	
3201	CAGCAACTGTGTAGAAGAGGAAGTATATCCAAACAACACA	3240
	TC C C G A A C C C	
3241	GTAACGTGTAATAATTATACTGGGACTCAAGAAGAATATG	3280
	T T C CG C C T A G G C	
3281	AGGGTACGTACACTTCTCGTAATCAAGGATATGACGAAGC	3320
	GA G C AGC CAG T CA	
3321	CTATGGTAATAACCCCTCCGTACCAGCTGATTACGCTTCA	3360
	TCC TCXXXXXXXXXX T T C T C C	
3361	GTCTATGAAGAAAAATCGTATACAGATGGACGAAGAGAGA	3400
	G C G G C C CA C T	
3401	ATCCTTGTGAATCTAACAGAGGCTATGGGGATTACACACC	3440
	C C G TC T CA C	
3441	ACTACCGGCTGGTTATGTAACAAAGGATTTAGAGTACTTC	3480
	T A T C T C GC T T	
3481	CCAGAGACCGATAAGGTATGGATTGAGATCGGAGAAACAG	3520
	T C A G C T C	
3521	AAGGAACATTCATCGTGGATAGCGTGGAATTACTCCTTAT	3560
	G C C GC T T G	
3561	GGAGGAA	3567

FIGURE 14E

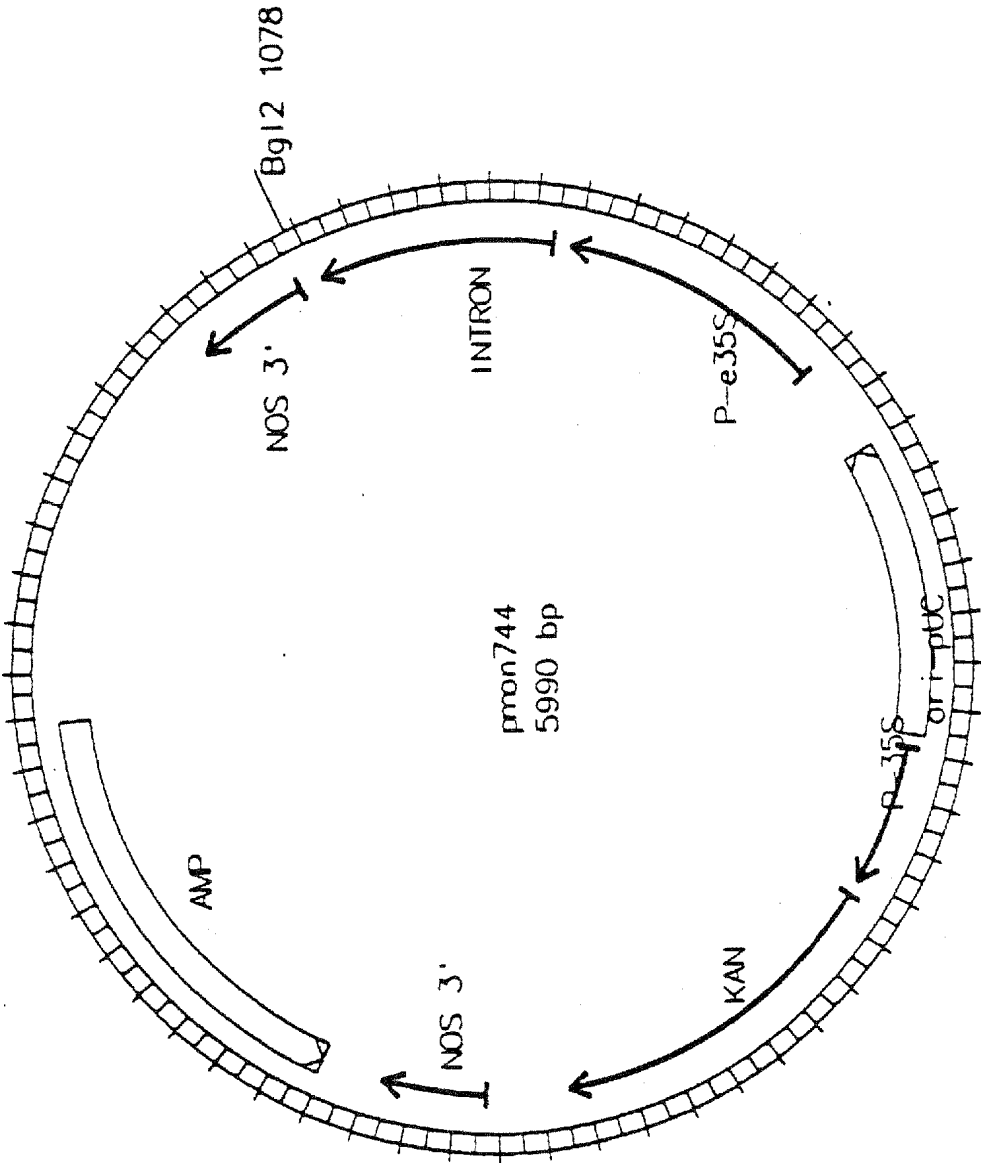


FIGURE 15

1	AGATCTAGAGGTAATTGTTATGAGTACTGTCGTGGTTAAG	40
	GATC	
41	GGAAACGTCAACGGTGGTGTACAACAACCTAGAAGGAGGA	80
	G T A	
81	GAAGGCAATCCCTTCGCAGGAGGGCTAACAGAGTACAGCC	120
	T A T	
121	AGTGGTTATGGTCACTGCTCCTGGCGAACCCAGGAGGAGG	160
	GC A A A	
161	AGACGCAGAAGAGGAGGCAATCGCAGGTCAAGAAGAAGT	200
	A G T A	
201	GAGTTCCCAGGGGAAGGGGCTCAAGCGAGACATTCGTGTT	240
	A A T	
241	TACAAAGGACAACCTCGTGGGCAACTCCCAAGGAAGTTTC	280
281	ACCTTCGGACCAAGTGTATCAGACTGTCCAGCATTCAAGG	320
	T	
321	ATGGAATACTCAAGGCCTACCATGAGTACAAGATCACAAG	360
	T	
361	TATCCTTCTTCAGTTCGTCAGCGAGGCCTCTTCCACCTCA	400
	T G T	
401	CCAGGATCCATCGCTTATGAGTTGGACCCACATTGCAAAG	440
	C A T	
441	TATCATCCCTCCAGTCCTACGTCAACAAGTTCCAAATCAC	480
	T	
481	AAAGGGAGGAGCTAAGACCTATCAAGCTAGGATGATCAAC	520
	T T C T	
521	GGAGTAGAATGGCACGATTCATCTGAGGATCAGTGCAGGA	560
	T T A	
561	TACTTTGGAAAGGAAGTGGAAAATCTTCAGACCCAGCAGG	600
	C A G T T	
601	ATCTTTCAGAGTCACCATCAGAGTGGCTCTTCAAACCCC	640
	T T A	
641	AAGTAATAGACTCCGGATCAGAGCCTGGTCCAAGCCCACA	680
	A T	

FIGURE 16A

681 ACCAACACCCACTCCAAC[.]TCCCCAAAGCATGAGCGATT[.]T 720
721 ATTGCTTACGTCGGCATACCTATGCTGACCATTCAAGAAT[.] 760
761 TC 762

FIGURE 16B